Journal of
Experimental Psychology

VOL. 27, NO. 5 NOVEMBER, 1940

A QUANTITATIVE STUDY OF COMBINATION TONES[1]

BY ERNEST GLEN WEVER, CHARLES W. BRAY, AND MERLE LAWRENCE

Princeton University

The combination tones were discovered nearly two centuries ago, and since that time many questions have arisen regarding their nature, origin, and relation to the more general problems of hearing. Subjective observations, on account of the complexities of the perceptual situation, have afforded only limited evidence as to the character of the phenomena and their dependence upon the stimuli. Such observations appear particularly deficient in attempts to determine the relative roles of peripheral and central portions of the auditory system.

The electrical responses of the ear have already been used in this field to demonstrate the general character of the distortion pattern and to reveal the parts of the auditory mechanism which are responsible for it (**16**, **17**, **18**). The present experiments are concerned with the combination tones and their quantitative relations to the stimuli, as these are shown in the electrical potentials of the cochlea.

METHOD AND PROCEDURE

The experiments were performed upon guinea pigs under Urethane anesthesia. The stimuli were pure tones from a loudspeaker, which was actuated by currents from two audio-frequency channels. Each channel consisted of a beat-fre-

[1] This investigation was aided by grants from the Rockefeller Foundation and the National Research Council.

quency oscillator, attenuators, filters, and other controls. The sounds from the loudspeaker were conducted through a tube to the meatus of the animal's ear. A probe tube with connected microphone was inserted in the system near the meatus for tests of purity. These tests showed that the sounds as delivered at the meatus contained only negligible amounts of harmonics and combination tones.

The cochlear potentials were picked up with a silver foil electrode on the membrane of the round window, and after amplification were studied with a wave analyzer (General Radio Type 736–A) and sometimes with a cathode ray oscillograph. Results were obtained on 20 ears, with a variety of procedures as described below.

For the discussion that follows, two definitions are necessary. By the *combination tone series* we refer to the complete series of components represented by the absolute values of $(mh \pm nl)$, when h, l are the primary frequencies and m, n take all possible integral values. The *order* of any combination tone is then $(m + n - 1)$.

The Combination-Tone Pattern

The first difference tone, which was the earliest combination tone to be found, is rather easily heard when the conditions are suitably arranged. Other combination tones, however, are heard with difficulty. Hällström (**3**) was able to observe about four difference tones. Helmholtz (**4**) noticed perhaps the same number, and in addition discovered the first summation tone. Later observers, with long practice and skillful arrangement of conditions, were able to detect components of fairly high orders, possibly as high as the 8th (**9**). However, in these situations only a few components, usually no more than two or three, could be made out at any one time. The exploring tone method, in which a third tone is brought in to beat with the component sought for, gives evidence that the auditory pattern is more extensive than it appears to direct perception. By this method Wegel and Lane (**15**) were able to demonstrate the existence of 11 or 12 combination tones, including members up to the 3rd order, when precautions were taken to insure purity of the stimuli.

In view of the limited indications of ordinary observations it is remarkable that the electrical responses of the cochlea reveal combination tones in veritable profusion. Without any attempt to exhaust the possibilities, we have recorded as many as 40 components, including overtones and combination tones, with stimuli of 1000~ and 2800~. These are shown in Table I, except for an 8th order tone of 6200~ which would require

TABLE I

COMBINATION TONES

(The arrangement shows the relation of the components to the primaries and their multiples.)

	h = 2800	$2h$ = 5600	$3h$ = 8400
l = 1000	1800 3800	4600 6600	7400 9400
$2l$ = 2000	800 4800	3600 7600	6400 10400
$3l$ = 3000	200 5800	2600 8600	5400 [11400]
$4l$ = 4000	1200 6800	1600 9600	4400 [12400]
$5l$ = 5000	2200 7800	600 10600	3400 [13400]
$6l$ = 6000	3200 8800	400 11600	2400 [14400]

another column. In the table, the numbers in brackets indicate frequencies looked for but not found. The 4 missing components are all in the last column, and all are summation tones; their high frequencies perhaps have something to do with their absence.

The highest order of component in the table is the 8th. With other primaries we have observed tones of yet higher order. Thus, with 1000~ and 1300~ we obtained 1800~, which is given both as $(7l - 4h)$ of the 10th order and as $(6h - 6l)$ of the 11th order. With 100~ and 10,000~ we obtained 8000~, which is $(h - 20l)$ and of the 20th order.

The production of combination tones does not seem to be limited by the frequency separation of the primaries. These

tones appear for primaries of closely similar frequency, and also for primaries from either end of the tonal scale. For example, we obtained a first difference tone of 25~ with 1000~ and 1025~; and with the primaries of 100~ and 10,000~ we found the complete series of difference tones $(h - l)$, $(h - 2l)$, $(h - 3l)$, etc. up to $(h - 20l)$, and the corresponding series of summation tones, $(h + l)$, $(h + 2l)$, $(h + 3l)$, etc. up to $(h + 15l)$, a total of 35 combination tones.

Newman, Stevens, and Davis (**10**) likewise found an extensive series of combination tones in the cochlear responses of the cat; they observed components up to the 11th order.

Components which do not belong to the harmonic series of either primary or to the combination tone series as defined by the formula $(mh \pm nl)$ have not been found. In addition to general exploration, we have made a special search for subharmonics, the meantone, and the intertone, but have not found them.

Intensity Relations

A number of systematic measurements have been made to discover the relation between the magnitude of the combination tones and the intensities of the two stimuli. Some of the results will first be presented for primaries of $h = 2800$~ and $l = 1000$~.

In the figures that follow, the magnitude of cochlear responses is shown in microvolts on the ordinate, and stimulus intensity is given in db. on the abscissa. The reference intensity (zero db.) is approximately 1 dyne per sq. cm.

Figure 1 shows the intensity relations for the first difference tone, $(h - l) = 1800$~. Three other curves appear in this figure, and are labelled 1000~, 2800~, and 1800~; these show the responses obtained from the two primary stimuli and from a stimulus tone of 1800~, each presented separately. The abscissa represents the intensity of the appropriate stimulus tone, for each of these curves. The three functions shown are simple power functions, as is general for the cochlear responses, and have slopes which vary only slightly from unity, except at high intensities where 'overloading' enters. The curve for 1800~ was not carried to the higher levels, as it was obtained

only for comparison with the difference tone of the same frequency.

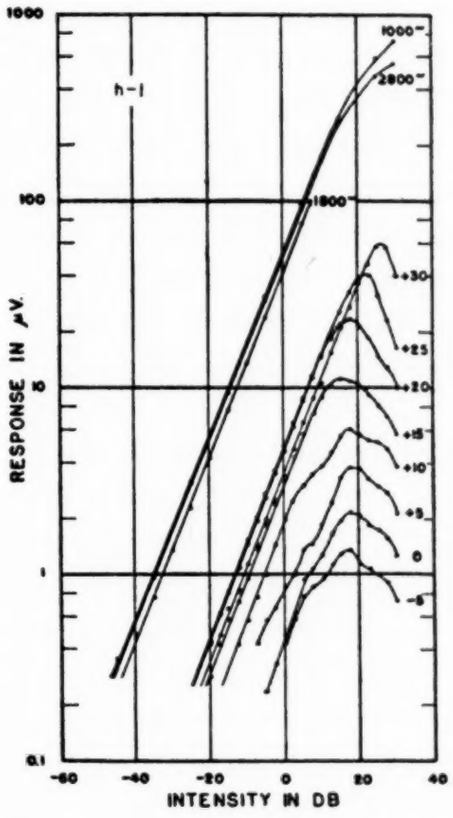


FIG. 1. Intensity functions for the first order difference tone $(h - l) = 1800\sim$, plotted with 2800~ as abscissa and 1000~ as parameter. Also, on the left, are functions for the primaries and a stimulus tone of 1800~.

The results for the difference tone form the family of curves on the right. They are plotted against the 2800~ stimulus as abscissa, and with the 1000~ stimulus as parameter. The parameter value (*i.e.*, the intensity of the 1000~ tone in db. with reference to 1 dyne per sq. cm.) is given near the end of each curve. The curve marked +30 db. thus represents the manner in which the first difference tone changes in magnitude when the 1000~ primary has an intensity of +30 db. and the 2800~ primary is varied from −20 to +30 db. A study of these results reveals the following relations:

(1) For any given intensity of the lower primary, the magnitude of $(h - l)$ rises regularly, and then passes through a maximum, as the intensity of the other primary is increased.

Beyond the maximum the rate of fall is sometimes rapid, sometimes rather irregular. For the higher values of the lower primary, the curves remain straight for a relatively large portion of their courses, and attain a higher maximum. For the lower values of this primary, the curves depart sooner from their straight courses, and show some irregularities.

(2) As we go from lower to higher parameter levels, the point of maximum remains for a time about the same abscissa value, and then shifts rather sharply to the right. It often appears that the shift begins about the level at which both primary functions depart from their simple form and begin to bend.

(3) Over their straight portions, before bending occurs, the $(h - l)$ curves are closely similar to the curves for the primaries, and also to the curves for the 1800~ stimulus. The steepness of these curves, as they appear on double-logarithmic coordinates, may be indicated by their tangents, or, more generally, as the value of a in the expression $E = kP^a$, where E = electrical potential in microvolts, P = sound pressure in dynes per sq. cm. for the stimulus that is varied, and k is a sensitivity constant. For the results shown, the values of the exponent a are 0.988 for the 1000~ curve, 0.986 for the 2800~ curve, 0.990 for the 1800~ stimulus curve, and they vary from 0.96 to 1.03 for the $(h - l)$ curves.

It must be emphasized that the exponents just given for $(h - l)$ are with respect only to the 2800~ stimulus, since for each curve the 1000~ primary is a constant. The contribution of the 1000~ primary thus appears in the formula only in k. If we replot the results with 1000~ as abscissa and 2800~ as parameter the result is a new family of curves closely similar to the one shown, with exponents which vary from 0.93 to 0.94. Indeed, the similarity of the two sets of curves is so close that it is unnecessary to present a new figure; and we need merely to state that the relations already described are true when either primary is the variable one.

At this point we raise the question whether the intensity functions of the combination tones are properly referred to their stimuli or to the cochlear functions obtained directly

from these stimuli. Since the primary functions are nearly linear with respect to their stimuli no decision can be made at the present time.[2]

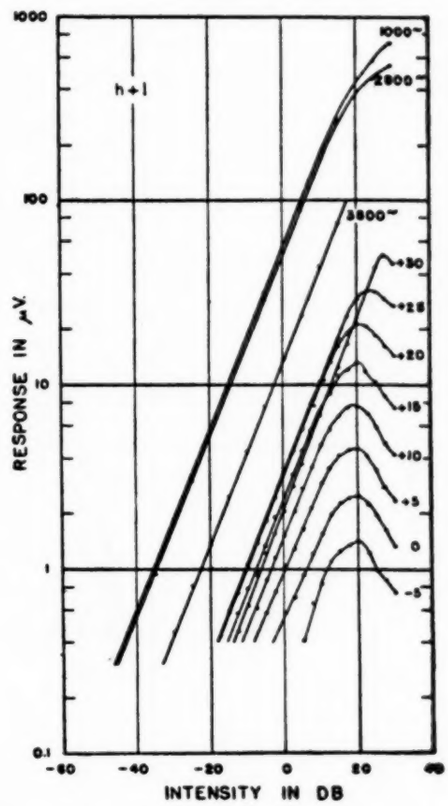


FIG. 2. Intensity functions for the first order summation tone $(h + l) = 3800\sim$ plotted as in Fig. 1.

Figure 2 presents results for the first summation tone, $(h + l) = 3800\sim$. As before, separate functions are shown for the two primaries and for a stimulus of 3800~. The curves for $(h + l)$ are plotted with 2800~ as abscissa and 1000~ as parameter. A similar family of curves would be obtained by plotting these results in the contrary manner.

In all essential respects the first summation tone shows the same stimulus relations as the first difference tone. The curves for the higher parameter values are simple throughout a considerable portion of their extent, and have exponents between 0.97 and 0.99 with respect to 2800~, as plotted; and between 0.98 and 1.06 with respect to 1000~. As before,

[2] An answer to this question will be sought in experiments with the pigeon, which gives a function which is not linear.

these exponents are similar to those for the primaries and for a stimulus tone of 3800~. The magnitudes of the summation tone curves are only slightly smaller than those of the corresponding difference tone curves.

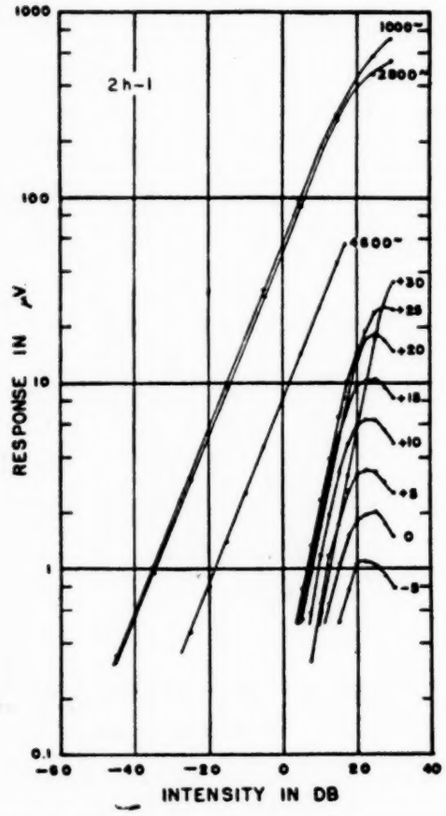


FIG. 3. Intensity functions for the second order difference tone $(2h - l) = 4600\sim$, plotted as in Fig. 1.

Figure 3 gives results for a difference tone of the second order, $(2h - l) = 4600\sim$, plotted with h as abscissa and l as parameter. Also shown are functions for the two primaries and a stimulus of 4600~. The curves for $(2h - l)$ are similar in many respects to those already given for the first order combination tones. They differ, however, in having higher slopes with respect to h, as plotted. The exponents for the uppermost curves vary from 1.82 to 1.98 as compared with the exponents already given for the primaries and an exponent of 0.990 for the 4600~ stimulus.

If these curves are replotted with l as abscissa and h as parameter they no longer assume the high slopes indicated,

but take on slopes closely similar to those shown in the two preceding figures. Exponents well above unity with respect to one or both of the primary stimuli are characteristic of the components of higher order, and will be treated systematically in the next section.

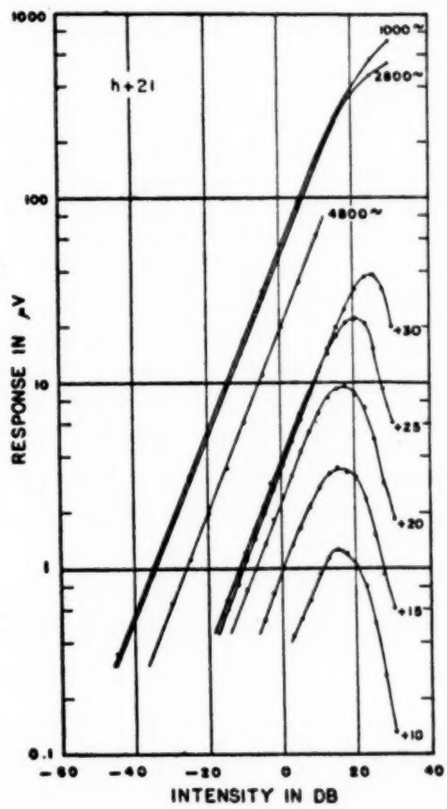


FIG. 4. Intensity functions for the second order summation tone $(h + 2l) = 4800\sim$, plotted as in Fig. 1.

Figure 4 shows the second order summation tone, $(h + 2l) = 4800\sim$, together with functions for the primaries and a stimulus of $4800\sim$. The $(h + 2l)$ curves are plotted with h as abscissa and l as parameter; and the uppermost curves have nearly linear slopes. If plotted in the contrary manner, with l as abscissa, their slopes would resemble those shown in Fig. 3.

The data just given are for a single pair of primaries, but many other pairs have been used with similar results. Fig. 5 shows the first difference tone as obtained with three pairs of primaries chosen from distinct regions of the frequency

scale. The primaries were 380~ and 4380~, 2300~ and 6300~, and 5200~ and 9200~, each pair of which gives 4000~ as $(h - l)$. The results for these three pairs of primaries are essentially alike. The curves for 380~ and 4380~ are not as

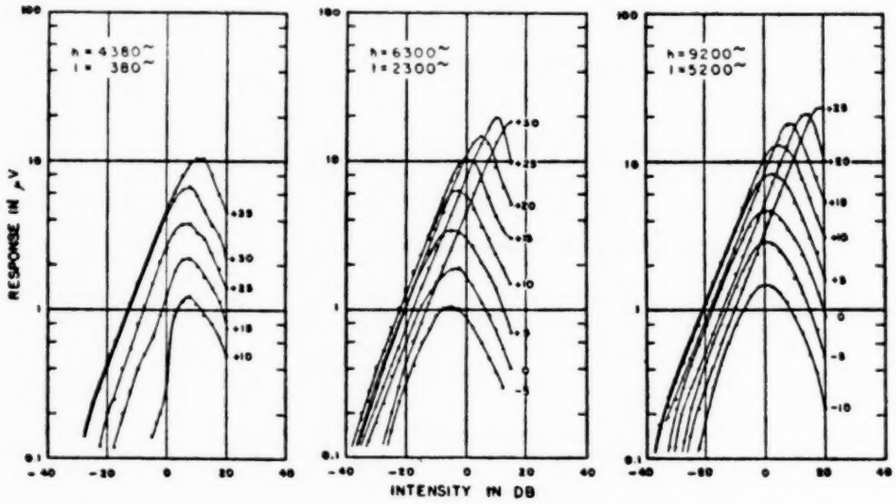


FIG. 5. First order difference tones obtained with primaries from different regions of the frequency scale, plotted against the higher primary in each case.

regular as the others, and for the stimulus intensities employed do not attain as high magnitudes, but otherwise their forms are much the same.

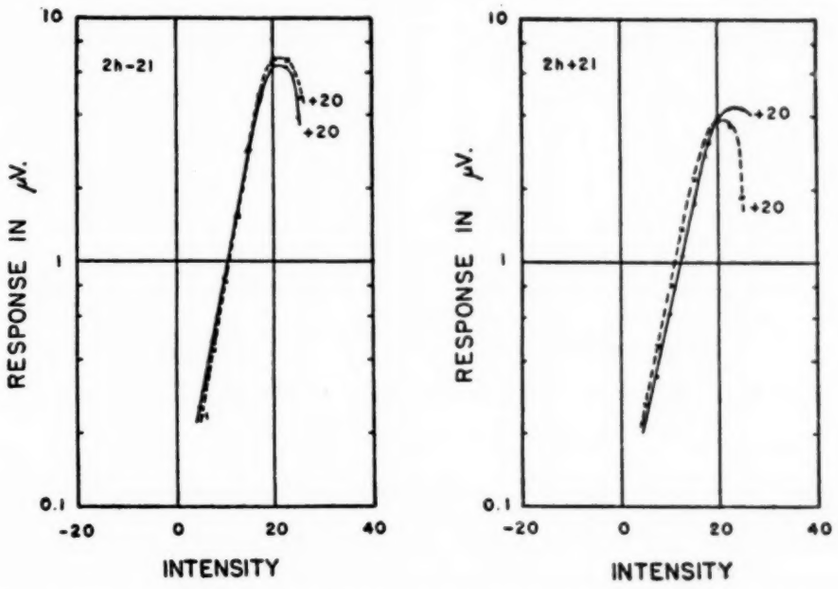
THE FUNCTIONS OF COMPONENTS OF HIGHER ORDER

The results of the preceding section have revealed the functional relation between the magnitude of the combination tone and the intensity of the two stimuli. Here we shall consider this relation further, with particular regard to the order of the component.

The intensity relation is simplest when one of the stimuli is strong and is held constant, while the other is varied over a wide range, for then the curve obtained is a power function throughout a large portion of its course. These functions have already been noted as the uppermost of the families of curves for the first two orders of combination tones. In Figs. 6 to 11 we present pairs of curves for combination tones of the 3rd, 4th, and 5th orders. For economy of space, the primary curves are not given. In other respects the arrange-

ment is similar to that of the preceding figures. The same primaries, $h = 3650\sim$ and $l = 1000\sim$, were used for Figs. 6–9.

Figure 6 shows results for the 3rd order difference tone $(2h - 2l) = 5300\sim$. The solid-lined curve shows this component as a function of h, with l constant at +20 db., and the dashed curve shows it as a function of l, with h constant at +20 db. The exponent of the function with respect to h is 2.08, and that with respect to l is 2.11.

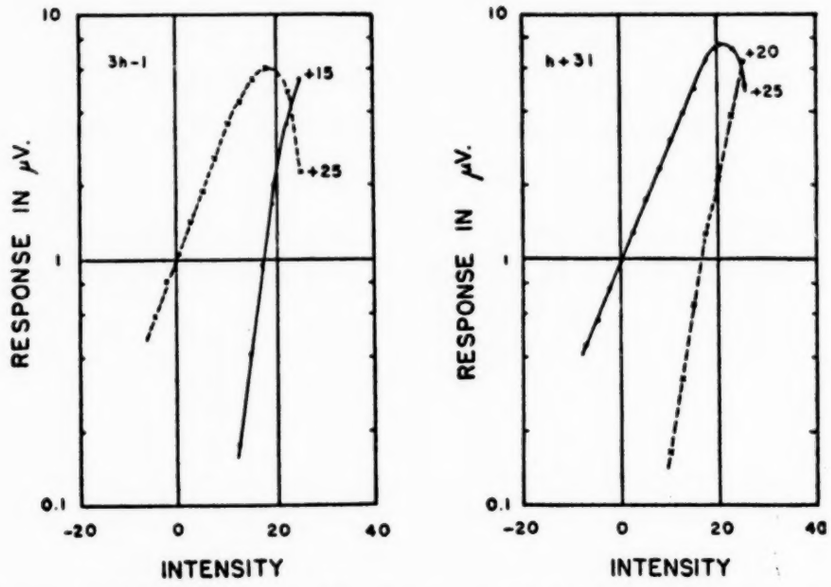


FIGS. 6 AND 7. Combination tones of the third order, of the form $(2h - 2l)$ and $(2h + 2l)$. The solid line shows the function with the higher primary, 3650~, as abscissa, and the dashed line with the lower primary, 1000~, as abscissa; in either case the parameter value is shown near the end of the curve.

Figure 7 shows results for the corresponding summation tone, $(2h + 2l) = 9300\sim$. Again, the solid-lined curve shows the component as a function of h, with l at +20 db., and the dashed curve shows it as a function of l, with h at +20 db. Here the exponents are 1.92 with respect to h and 1.93 with respect to l.

Figure 8 presents curves for a 3d order difference tone of different composition, $(3h - l) = 9950\sim$. The unbroken curve is plotted with respect to h, with l at +15 db., and the dashed curve with respect to l, with h at +25 db. The marked difference in steepness of the two functions is apparent. The function with respect to h has an exponent of 3.06, and that with respect to l has one of 1.02.

Figure 9 shows results for the summation tone $(h + 3l) = 6650\sim$. The solid line is plotted with respect to h with l at +25 db., and the dashed line with respect to l with h at



FIGS. 8 AND 9. Third order combination tones of the form $(3h - l)$ and $(h + 3l)$, plotted as in Figs. 6 and 7.

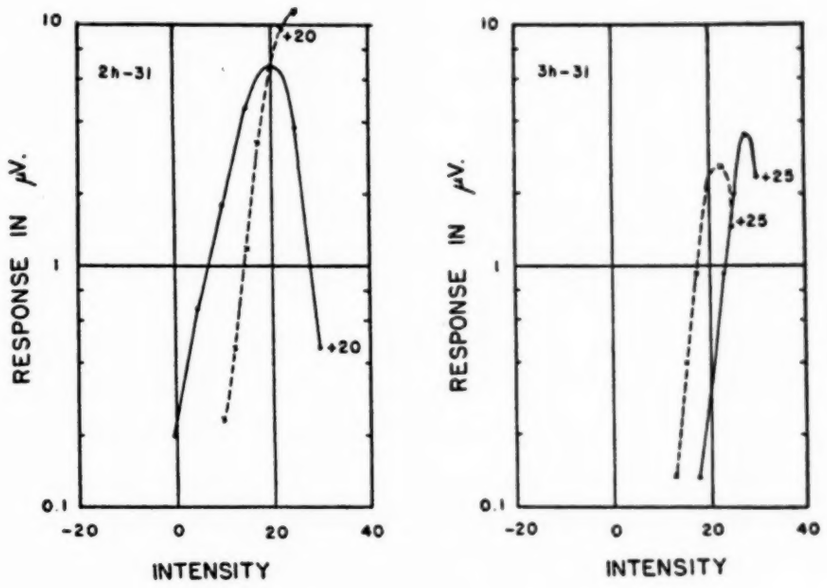
+20 db. Again, there is a difference in steepness, but it will be noted that here the roles of the primaries are reversed as compared with the preceding example. The exponent for h is 0.95, and that for l is 2.57.

Figures 10 and 11 give results on difference tones of the 4th and 5th orders, with different primaries. Fig. 10 presents curves for the 4th order component $(2h - 3l) = 4700\sim$, produced by primaries of $3800\sim$ and $966\frac{2}{3}\sim$. The solid line is plotted with respect to h, with l at +20 db., and the dashed line with respect to l, with h at +20 db. The exponent for h is 1.81, and that for l is 3.13.

Figure 11 gives results for a difference tone of the 5th order, $(3h - 3l) = 4700\sim$, obtained with primaries of $2533\frac{1}{3}\sim$ and $966\frac{2}{3}\sim$. The solid line is plotted with respect to h, with l at +25 db., and the dashed line with respect to l, with h at +25 db. The exponents are 3.09 for h and 3.37 for l.

It is clear from the foregoing evidence that the order of the function as represented by the exponent a is a significant feature in the relation of the combination tone to its stimuli,

within the limits where the function is a simple one. The value of the exponent varies systematically with the nature of the component, or, more exactly, with the manner in which



FIGS. 10 AND 11. Combination tones of the fourth and fifth orders. On the eft is $(2h - 3l)$, and on the right $(3h - 3l)$. As in the preceding figures, the solid line shows the function with respect to the higher primary, and the dashed line shows the converse function.

the variable stimulus enters into the formation of the component. A comprehensive study was made of the values of a for several orders of combination tones, as given below.

Series 1.—Five pairs of primaries were selected to give successive orders of difference tones, all with a frequency of 1800~. They were as follows:

D_1 : $h =$ 2800~, $l =$ 1000~; $(h - l) =$ 1800~.
D_2 : $h =$ 3800~, $l =$ 2800~; $(2l - h) =$ 1800~.
D_3 : $h =$ 6600~, $l =$ 2800~; $(3l = h) =$ 1800~.
D_4 : $h =$ 9400~, $l =$ 2800~; $(4l - h) =$ 1800~.
D_5 : $h =$ 12200~, $l =$ 2800~; $(5l - h) =$ 1800~.

The functional relations to stimulus intensity were studied not only for the difference tones but also for the primaries and for certain harmonics of the lower primary. The exponents of the functions are shown in Table II.

TABLE II

VALUES OF *a* FOR FIVE ORDERS OF DIFFERENCE TONES

	D_1	D_2	D_3	D_4	D_5
f(h)	0.98	0.95	0.99	1.09	0.99
f(l)	1.00	1.01	1.00	1.02	1.02
$f_h(D)$	1.02	1.02	0.95	1.00	1.67
$f_l(D)$	1.08	2.10	2.98	3.07	2.42
$f_l(nl)$	—	2.11	2.82	3.07	3.76

In the table, the data for the several orders of difference tones are shown in the columns headed by the designations D_1, D_2, etc. The rows marked f(h) and f(l) show the exponents of the primaries themselves. The row marked $f_h(D)$ shows the exponents for the difference tones as a function of h, when l is constant, and, correspondingly, the row marked $f_l(D)$ shows these data when l is the variable stimulus and h is constant. The last row shows the exponents of the functions for harmonics of l; here n expresses the order of the harmonic, which for any given column is the same as the order of the indicated difference tone. For example, the column headed D_3 shows that from primaries with exponents of 0.99 and 1.00 was obtained a 3d order difference tone ($3l - h$) whose exponent with respect to h was 0.95 and with respect to l was 2.98; also the 3d harmonic of the lower primary, $3l$ = 8400~, had an exponent of 2.82.

The measurements on the difference tones beyond the 3d order were scanty, as under the conditions these components reached levels barely within the range of sensitivity of our apparatus. The curves resembled the lowermost curves in Figs. 1–5, and thus represented the function only in the region of the maximum. Therefore our measurements of slope for these components are not very significant. For the components of the three lower orders, however, certain relations hold. As indicated in the table, exponents for the two primaries and for the difference tones as a function of h are of the same order of magnitude and approximate unity. The

exponents for the difference tones as a function of l increase with the order of the difference tone. Similarly, the exponents for the harmonics of l increase with the order of the harmonic.

Series 2.—Stimuli were chosen to give the first three orders of summation tones, corresponding to the series of difference tones just described. Also, other pairs of primaries were added to give 3d and 4th order summation tones in which both h and l entered in complex relationship, as shown. All these summation tones had the same frequency.

S_1 : $h = 2900\sim$, $l = 1800\sim$; $(h + l) = 4700\sim$.
S_2 : $h = 2900\sim$, $l = 900\sim$; $(h + 2l) = 4700\sim$.
S_{3a}: $h = 2900\sim$, $l = 600\sim$; $(h + 3l) = 4700\sim$.
S_{3b}: $h = 1450\sim$, $l = 900\sim$; $(2h + 2l) = 4700\sim$.
S_{4a}: $h = 966\frac{2}{3}\sim$, $l = 900\sim$; $(3h + 2l) = 4700\sim$.
S_{4b}: $h = 1450\sim$, $l = 600\sim$; $(2h + 3l) = 4700\sim$.

TABLE III

VALUES OF a FOR SUMMATION TONES

	S_1	S_2	S_{3a}	S_{3b}	S_{4a}	S_{4b}
$f_h(S)$	1.07	1.09	1.16	1.98	2.94	1.69
$f_l(S)$	1.02	2.04	2.79	1.88	2.28	2.93

In Table III, values of the exponents are given for the summation tones as a function of the two primary stimuli. These results show clearly that the size of the exponent with respect to a given primary is determined by the manner in which that primary enters into the formation of the combination tone. The higher the order of the primary as it enters into the combination tone, the larger is the exponent with respect to that primary. This relation will be stated more precisely in algebraic terms, as follows:

$$f(h \pm l) \propto P_h P_l$$
$$f(2l \pm h) \propto P_h P_l^2$$
$$f(3l \pm h) \propto P_h P_l^3$$

or, for the general case,

$$f(mh \pm nl) \propto P_h^m P_l^n,$$

in which the symbol f indicates the intensity function of the tone in parentheses, P_h, P_l designate the stimulus intensities of the two primaries, and m, n are any integers. The first expression thus signifies that the magnitude of the first difference or summation tone varies directly as the product of the intensities of the two stimuli.

The primary functions themselves are $f(h) \propto P_h$ and $f(l) \propto P_l$; and the functions for the overtones, according to Table II, have the form $f_h(mh) \propto P_h^m$ and $f_l(nl) \propto P_l^n$.

These statements are true only if we 'round off' the obtained values of a to their nearest integral values. This procedure is justified only if the variations are due to experimental error. We can obtain some evidence whether this is true by statistical treatment of a large number of measurements. Table IV gives average results of determinations of

TABLE IV

MEAN VALUES OF a FOR DIFFERENT VALUES OF m

	$m = 1$	$m = 2$	$m = 3$
Mean	1.02	2.06	2.83
Mean variation	0.08	0.09	0.36
Range	0.85–1.59	1.57–2.86	1.83–3.62
Number of cases	70	50	23

exponents for a large number of combination tones. In the table, the data have been grouped without regard to the order of the combination tone itself, or to whether h or l was the variable stimulus. By $m = 1$, 2, or 3 is indicated the order in which a given primary entered into the combination tone. Each combination tone thus usually yielded two data for the table: *e.g.*, for the component $(h + 3l)$, the variation with respect to h appears under $m = 1$, and the variation with respect to l under $m = 3$. The table gives, for the three orders of m, the mean value of exponent a obtained, the mean variations, the ranges, and the number of measurements.

As the table shows, the mean values of the exponents for $m = 1$ and $m = 2$ are very nearly integral, though the range of variation is large. For $m = 3$ the mean is below 3, but the difference is small, and under the conditions is probably negligible.

Theoretical Considerations

Three principal theories have been advanced to account for the combination tones. These are the beat-tone theory, the overtone theory, and the transformation theory.

The Beat-Tone Theory.—Young (**19**) first suggested that when beats become sufficiently rapid they pass over into a tone; this tone is what we now call the first difference tone. Hällström (**3**) developed the hypothesis to account for the difference tones of higher orders. His assumption was that a difference tone already present might interact with one of the primaries to produce other beats, and through these bring about additional difference tones. The scheme is as follows, with examples for particular primaries:

$$D_1 = (h - l). \quad (2800 - 1000 = 1800)$$
$$D_2 = (D_1 - l). \quad (1800 - 1000 = 800)$$
$$D_3 = (l - D_2). \quad (1000 - 800 = 200)$$
etc.

One form of the beat-tone theory has achieved practically independent status as the *resultant displacements theory.* This theory accounts for the combination tones on the assumption that any periodic variation in the form of a complex wave can be appreciated as a tone. The theory is subject to the serious criticism that no way has yet been shown by which a phase or amplitude variation can function as a frequency variation.

The Overtone Theory.—Ohm (**11**) objected to the secondary interaction of difference tones and primaries, which was the basis of Hällström's theory, and proposed instead that the difference tones of higher order result from overtones. He believed that the overtones were in the stimuli, which would make all the difference tones be of the first order as defined by Hällström. The three components just mentioned were thus obtained as

$$D_1 = (\ h - l). \quad (2800 - 1000 = 1800)$$
$$D_2 = (h - 2l). \quad (2800 - 2000 = 800)$$
$$D_3 = (3l - h). \quad (3000 - 2800 = 200).$$

It was later shown by Koenig (**8**) and others that the combination tones continue to be heard even when the primaries are carefully freed of overtones. It was then presumed that the overtones could arise within the ear itself, though no clear explanation was offered for the process apart from the transformation theory, next to be discussed.

The Transformation Theory.—Helmholtz (**4**, **5**) postulated that the combination tones, and overtones as well, arise through distortion in the middle ear. Specifically, he suggested that displacements as represented at the inner ear are proportional not simply to the impressed force, but also in significant degree to the square of the force, or

$$X = a_0 + a_1P + a_2P^2,$$

where X is the displacement, P is force per unit area, or sound pressure, and a_0, a_1, a_2 are constants. Due to the presence of the square term in the equation, its expansion for a single sinusoidal wave gives a term of double frequency as well as the impressed frequency itself; and an expansion for two waves, when P becomes the sum of the individual pressures, gives sum and difference frequencies in addition to the primaries and their overtones.

Helmholtz appreciated that even higher components enter if X becomes sensibly proportional to higher powers of P. Others have carried out the necessary computations. The most extensive treatment is that of Cotton (**2**), who expanded the formula to the 7th power of P, and thus included all combination tones up to the 6th order.

For the complete formula the reader is referred to Cotton's paper. Here it will suffice, for further reference, to list the first 4 terms, with the last 3 in abbreviated form. The component is given first, and then in brackets its coefficient. Here ω_h, ω_l correspond to the primaries h and l (*e.g.*, $\omega_h = 2\pi h$); P_h, P_l indicate their respective pressures; and a_1, a_2, a_3 . . . are arbitrary constants from the power series. In reproducing

these expressions we have altered the subscripts to conform with practice elsewhere in this paper.

$$(1)\ \cos(\omega_h \pm \omega_l)t[a_2P_hP_l + 1.5a_4(P_h^3P_l + P_hP_l^3) + 1.87a_6(P_hP_l^5 + 3P_h^3P_l^3 + P_h^5P_l) + \ldots]$$

$$(2)\ \cos(2\omega_h \pm \omega_l)t[0.75a_3P_h^2P_l + * * *]$$

$$(3)\ \cos(\omega_h \pm 2\omega_l)t[0.75a_3P_hP_l^2 + * * *]$$

$$(4)\ \cos(3\omega_h \pm \omega_l)t[0.5a_4P_h^3P_l + * * *]$$

Other terms, including the primaries, their series of overtones, and a constant displacement term, are not given.

Both the beat-tone theory and the overtone theory are defective, in that they fail to account for all the combination tones that appear in Table I. Neither theory accounts for the summation tones, while the beat-tone theory omits a large number of the difference tones as well. This objection to the theories of course is not new, as it emerged at once on Helmholtz's discovery of the summation tones. The reply made by the supporters of the theories, and in particular by upholders of the overtone theory, was that the summation tones did not exist, or that they actually were difference tones of higher order. We shall examine the latter contention.

The Status of the Summation Tone.—The assertion that the summation tone is a difference tone of higher order rests upon the fact that we can always derive a difference tone of the necessary frequency by assumption of the interaction of overtones. It is easily shown, however, that the necessary overtones are oftentimes of very high orders. For the primaries most frequently employed in this study, 2800~ and 1000~, the frequency of 3800~, which is that of the first summation tone, can be obtained as a difference tone if we assume the presence of the 6th harmonic of one primary and the 13th harmonic of the other, for if these are present we can obtain $(6h - 13l) = 3800$~. For some other ratio of the primaries quite different overtones are required, and this fact in itself is a defect of the theory.

Some time ago, Peterson (**12**) sought evidence whether an observed component was a summation tone or a difference tone of higher order by the method of mistuning one of the

primaries. The argument is that mistuning must have a different effect on the two possible components. To illustrate, in the observation of $(6h - 13l)$, an increase of l by $1^{\sim}$ should reduce the difference frequency by $13^{\sim}$; whereas if the component is $(h + l)$ such a change should increase the frequency by $1^{\sim}$.

We have observed in this situation, as Peterson did, that a change of l is accompanied by a change of the combination tone in equal amount and in the same direction. The component therefore behaves like $(h + l)$, and not like a difference tone of higher order like $(6h - 13l)$.

To this argument it may be objected that a change in the frequency ratio of the primaries alters the particular overtones required, as mentioned above. However, a change of the overtones in action would alter the frequency of the component most drastically, and could never account for the regular and progressive alteration observed.

One further contention needs to be considered. Preyer (**13**) accepted the view that the first summation tone was actually a product of the interaction of the first difference tone and the second harmonic of the higher tone, *i.e.*, was $[2h - (h - l)]$. Since algebraically this quantity reduces to $(h + l)$, it must react to mistuning in the same manner; and the above observations on the effects of mistuning are not discriminative. However, Preyer's explanation of the tone in question can be objected to on two grounds. (1) It is illogical to suppose that the first difference tone itself can enter secondarily into interaction. The reply that has sometimes been made to escape this criticism, that there are two successive primary interactions in different places, *i.e.*, one in the middle ear that gives $2h$ and $(h - l)$, and then another in the inner ear that gives $[2h - (h - l)]$, is fairly well countered by the evidence that in the absence of the middle ear (up to the stapes) the summation tones nevertheless appear (**18**). (2) Direct evidence is supplied by observations of the variation in magnitude of the component with respect to the intensity of h. If the component were $[2h - (h - l)]$ it should vary as P_h^3; whereas actually it varies as P_h. Hence it is concluded that the component is $(h + l)$.

If any further evidence is needed on the status of the summation tone it is afforded by the following observations, already reported:

1. The summation tone is of the same general magnitude as the corresponding difference tone. Compare, for example, Figs. 1 and 2, or 6 and 7. Nearly always, for particular stimuli, the difference tone slightly exceeds the summation tone of corresponding composition, but occasionally the reverse is true. These differences may be due in part to sensitivity factors.

2. The summation tones are almost as numerous as the difference tones, and parallel them closely. These facts are evident in Table I; the only variations appear for the higher orders where the summation tones reach extreme frequencies.

3. Summation tones and difference tones of the same orders have similar functions with respect to their stimuli. This fact is clearly evident in the figures and especially in the values of the exponents presented in Tables II and III.

The above evidence disposes of the first two theories, at least in their conventional forms, and we are left with the transformation theory.

Evaluation of the Transformation Theory.—Before entering into a consideration of the transformation theory in the light of the evidence now available it is necessary to point out that in principle the theory remains valid even if we reject the original view of Helmholtz regarding the seat of the distortion. He attributed the transformation process to the middle ear. Recent evidence, however, shows that distortion occurs in the inner ear, and, under most conditions at least, this inner ear distortion is considerably more important than any that may occur in the more peripheral mechanism (**17**, **18**). We speak, therefore, of the transformation theory in a general sense.

The transformation theory satisfies the initial requirements of the situation by accounting for all demonstrable components, combination tones and overtones alike. To do so, however, it is necessary to postulate a power series of very great length, beyond any present mathematical derivations. Cotton, as just mentioned, developed the formula through the

7th power of P, and obtained the coefficients of the first 42 combination tones. Even this is insufficient, as our observation of components as high as $(h - 20l)$ indicates that the expansion must be carried at least as far as the 21st power. Such an expansion would give 420 combination tones.

The transformation theory is in good agreement with the evidence in placing the difference tones and the summation tones on the same footing. In the expansion of the fundamental formula these components come out in equal numbers, and those of similar order and form have identical coefficients. The observed variations in the relative magnitude of difference and summation tones are small, and suggest no serious fault in the theory. As already indicated, they probably reflect the presence of sensitivity factors in the ear, which enter either coincidently with the transformation process or following it.

When we consider the variation of the combination tones with respect to stimulus intensity it becomes necessary to modify the formula somewhat. According to the mathematical derivations, the components are contributed to by various (alternate) powers of P. Thus, the combination tones of the first order are determined not only by P^2, but also by P^4, P^6, . . . , while those of the 2d order are determined by P^3, P^5, ... Let us consider only the components of the first order, for which we have given the whole coefficient as developed by Cotton. Now if in this coefficient the constants a_4, a_6, . . . are appreciable with respect to a_2, the function for the component will rise relatively rapidly with respect to either primary. For example, with respect to h, the function is in the first power to an extent determined by a_2, in the 3d power to an extent determined by a_4 and a_6, and in the 5th power to an extent determined by a_6; at the same time these higher powers contribute considerably to the first power variation. Since we find empirically that the function varies simply in the first power with respect to h, it is evident that the terms of the coefficient beyond the first must be negligible. Yet we cannot make them so by assuming that the constants a_1, a_6 are negligible, as higher order components containing

these constants are found in considerable magnitude. It appears that the procedure which forms the basis of the mathematical derivations needs to be reëxamined.

If we neglect all but the first term of each coefficient we find excellent conformity with the observations. The exponents given for P_h and P_l, which express the rate of variation of the combination tone with the intensities of the primary stimuli, agree within reasonable limits with our observed values, at least for the lower orders for which our data are sufficient. We therefore revise the coefficients for the combination tones, as follows:

(1) $\cos(\omega_h \pm \omega_l)t[a_2 P_h P_l]$
(2) $\cos(2\omega_h \pm \omega_l)t[a_3 P_h^2 P_l]$
(3) $\cos(\omega_h \pm 2\omega_l)t[a_3 P_h P_l^2]$
(4) $\cos(3\omega_h \pm \omega_l)t[a_4 P_h^3 P_l]$

or, in general,

$$\cos(m\omega_h \pm n\omega_l)t[a_{m+n} P_h^m P_l^n].$$

This is, of course, merely the trigonometric form of the expression already given as

$$f(mh \pm nl) \propto P_h^m P_l^n.$$

If we take account of sensitivity factors, we must alter the formulas by the addition of new constants b_2, b_3, . . . b_{m+n}, which probably vary somewhat for each component. The general formula then will read

$$\cos(m\omega_h \pm n\omega_l)t[a_{m+n} b_{m+n} P_h^m P_l^n].$$

Perceptual Relations

The results of this investigation reveal the presence and character of the combination tones in the peripheral processes of the ear. We now consider the relation of these processes to the phenomena shown in perception.

As already mentioned, direct observation of the combination tones is a difficult matter. It has long been known that the observation is facilitated by proper manipulation of the intensities of the stimuli.

It was asserted by Hermann (7) that no advantage is gained in the hearing of a difference tone if one primary is raised in intensity while the other is held constant. Indeed, according to his account, this procedure is disadvantageous, for the difference tone, though audible for equal and rather weak primaries, promptly disappeared when one was increased in strength. The results of the present study show that in the peripheral processes the combination tone increases when, with one primary constant, the other is raised from a low level of intensity. However, at a particular point, indicated as the maximum of the curve, this increase ceases, and a further rise of the primary intensity impairs the combination tone in the manner described by Hermann. It is likely that Hermann's choice of intensities was such that he was working only with the latter part of the intensity function, from the maximum onward.

Helmholtz (5) noted that a combination tone is raised in strength both absolutely and relatively when both primaries are simultaneously increased. This observation is supported by our results. It may be seen in the formula given above that a combination tone is proportional to the product of the two primary intensities (each raised to its proper power). If such is the case it is evident that the combination tone must rise more rapidly than the sum of the primary intensities; and the rate of rise is greater the higher the order of the combination tone.

As has already been made plain, the relation just described is valid only within the limits where a power function obtains. If the intensity is carried sufficiently far this function fails, the curve approaches a maximum, and finally bends downward. When one primary is constant, there is a particular intensity of the other that gives optimum conditions for observation of any combination tone. When both primaries are varied the optimum conditions are probably beyond any intensity that the ear can withstand without damage.

It has occasionally been reported that the second order difference tone is heard more readily than the first order difference tone. Our results show clearly why this relation may

occur. Though the second order tone usually begins to be detectable at a higher stimulus level than the first order tone, it rises more rapidly, and sometimes it attains a higher maximum. This condition is clearly shown in curves previously published, which were obtained by simultaneous variation of the two stimuli (**18**).

The Role of Masking.—Previous observers have reported a particular difficulty in hearing what has been named the intervening difference tone, which is a difference tone whose frequency lies between the two primary frequencies. The first difference tone has this position when the ratio of the primaries is greater than an octave. The primaries most used in the present study, 1000~ and 2800~, give this kind of first difference tone (1800~). Schaefer (**14**) went so far as to suggest that the intervening difference tone did not exist, or else was so faint as to be inaudible under most conditions. Several other observers, however, were able to convince themselves of its presence.

A combination tone that falls outside the range of the primaries, and particularly one that falls well below them, is in a better position to be heard than one of the same order that is in the intervening position. Indeed, the early observers referred to the combination tones as 'grave harmonics,' because they noticed only the low-frequency tones.

The reason for the differences in audibility of these types of combination tones is not found in the peripheral processes, and must arise in processes that follow. The most reasonable explanation is the differential effect of masking. The difference tone that falls between strong primaries will be in the masking area of each of them, and will be largely or wholly suppressed. One that falls above the primaries will be affected less, though it may be somewhat impaired by the harmonic masking areas. One that comes well below the primaries will be affected little or not at all.

The summation tones as a class have always been more difficult to observe than the difference tones. Some of the earlier writers were never convinced of their existence. Hermann (**6**) once remarked that he himself "had never been

able to hear them, and he could not find anyone who had." Barton (1) attempted a mathematical explanation of the feebleness of these tones in comparison with the difference tones. According to his computations, the difference tone should be inversely proportional to the square of the difference in the primary frequencies, and the summation tone inversely proportional to the square of their sum. It is evident from results already presented that this relation does not hold for the peripheral processes.

The perceptual faintness of the summation tones can be accounted for by masking. These tones are necessarily of higher frequency than the primaries, and can hardly fail to be partially masked by the primaries themselves or their harmonics.

The above considerations make clear why manipulation of the primary frequencies can bring about more favorable conditions for the observation of particular types of combination tones, as the proper adjustments will move the desired component out of the masking areas of the primaries. This point can more easily be appreciated if we consider the forms of masking curves for strong stimuli (15). When two stimuli are present the resulting masking areas will be extensive and complex, and it is likely that only a few regions of frequency will be free of their effects. The older investigators evidently attained considerable skill in manipulating the primary ratios for favorable observations, even though they proceeded for the most part by trial, and not with any guiding principles.

Summary

The electrical responses of the cochlea were used to study the combination tones and their quantitative relations to the stimuli.

1. An analysis of the response to two pure tones of high intensity revealed, in addition to the harmonics of the two primaries, all the sum and difference combinations of the stimuli and their integral multiples, up to rather high orders. Other components, such as subharmonics, the meantone, and the intertone, were not found.

2. The magnitude of any combination tone is a complex function of the intensity of the two stimuli. When one stimulus is strong and is held constant while the other is varied over a wide range, the combination tone follows a power function for a considerable portion of its course, then finally, when the variable stimulus becomes relatively strong, it passes through a maximum and bends downward.

3. Within the limits where the combination tone follows a power function there is a systematic relation between the rate of variation of the magnitude of the combination tone with respect to a given stimulus and the manner in which that stimulus enters into the formation of the combination tone. At least for the combination tones of lower orders, which are large enough for precise measurements, this relation can be given a precise mathematical formulation as $f(mh \pm nl) \propto P_h{}^m P_l{}^n$.

4. The results are considered in relation to the classical theories on the origin of combination tones: the beat-tone theory, the overtone theory, and the transformation theory. Of these, the transformation theory is found to be correct in principle, but in need of correction in mathematical detail.

5. In the response pattern, the summation tones have equal status with the difference tones; they cannot be accounted for as difference tones of higher order. For particular primaries there are only slight differences in magnitude between the summation tones and their corresponding difference tones; usually the difference tones are slightly larger, but occasionally the reverse is true.

6. A number of perceptual phenomena, heretofore obscure, receive a simple explanation by way of the peripheral processes. The results show why it is usually an advantage in hearing the combination tones to raise the intensity of both stimuli; and also why it may be a disadvantage, under some conditions, to raise but one of them.

7. Some features of the perceptual phenomena are not found in the peripheral processes, and must be ascribed to supervening activities in the nervous system. Among such activities is masking, which is called upon to account for the

relative difficulty of hearing the intervening difference tone and the summation tones.

(Manuscript received June 10, 1940)

References

1. Barton, E. H., *A text-book on sound*, London: Macmillan, 1922, *pp.* 171 *ff.*, 390 *ff.*
2. Cotton, J. C., Beats and combination tones at intervals between the unison and the octave, *J. Acoust. Soc. Amer.*, 1935, **7**, 44–50.
3. Hällström, G. G., Von den Combinationstönen, *Ann. Phys.*, 2nd series, 1832, **24**, 438–466.
4. Helmholtz, H. L. F., Ueber Combinationstöne, *Ann. Phys.*, 2nd series, 1856, **99**, 497–540.
5. Helmholtz, H. L. F., *Sensations of tone*, Trans. by A. J. Ellis, London: Longmans (2nd Eng. Ed.), 1885, *pp.* 152 *ff.*, 411 *ff.*
6. Hermann, L., Zur Theorie der Combinationstöne, *Pflüg. Arch. Physiol.*, 1891, **49**, 499–518.
7. Hermann, L., Neue Untersuchungen über die Natur der Kombinationstöne, *Pflüg. Arch. Physiol.*, 1908, **122**, 419–447.
8. Koenig, R., Ueber den Zusammenklang zweier Töne, *Ann. Phys.*, 2nd series, 1876, **157**, 177–237.
9. Krüger, F., Zur Theorie der Combinationstöne, *Phil. Stud.*, 1901, **17**, 185–310.
10. Newman, E. B., Stevens, S. S., and Davis, H., Factors in the production of aural harmonics and combination tones, *J. Acoust. Soc. Amer.*, 1937, **9**, 107–118.
11. Ohm, G. S., Bemerkungen über Combinationstöne und Stösse, *Ann. Phys.*, 2nd series, 1839, **47**, 463–466.
12. Peterson, J., Origin of higher orders of combination tones, *Psychol. Rev.*, 1915, **22**, 512–518.
13. Preyer, W., *Akustische Untersuchungen*, Jena: Fischer, 1879, p. 12.
14. Schaefer, K. L., Eine neue Erklärung der subjectiven Combinationstöne auf Grund der Helmholtz'schen Resonanzhypothese, *Pflüg. Arch. Physiol.*, 1899, **78**, 505–526.
15. Wegel, R. L., and Lane, C. E., The auditory masking of one pure tone by another, etc., *Phys. Rev.*, 1924, **23**, 266–285.
16. Wever, E. G., and Bray, C. W., Distortion in the ear as shown by the electrical responses of the cochlea, *J. Acoust. Soc. Amer.*, 1938, **9**, 227–233.
17. Wever, E. G., Bray, C. W., and Lawrence, M., The locus of distortion in the ear, *J. Acoust. Soc. Amer.*, 1940, **11**, 427–433.
18. Wever, E. G., Bray, C. W., and Lawrence, M., The origin of combination tones, *J. Exp. Psychol.*, 1940, **27**, 217–226.
19. Young, T., Outlines of experiments and inquiries respecting sound and light, *Phil. Trans. Roy. Soc. Lond.*, 1800, Part 1, 130–133.

ANXIETY-REDUCTION AND LEARNING

BY O. H. MOWRER *

Department of Psychology, Institute of Human Relations, Yale University

The Problem

That a reduction in the intensity of the psychosomatic state known as dread, apprehension, nervousness, tension, fear, or anxiety, can function as a 'satisfying state of affairs' (as defined by Thorndike's law of effect) and thereby bring about the learning of new stimulus-response sequences ('connections') is indicated by various clinical as well as commonsense observations (**8**, **9**, **16**). So far as the writer is aware, however, no previous attempt has been made to submit this proposition to experimental verification.[1] More specifically, the hypothesis under examination here is that just as a reduction in hunger, reduction in thirst, reduction in sex drive, reduction in fatigue, reduction in oxygen lack, or reduction in any other organic 'need' or 'discomfort' tends to 'reinforce' the behavior which brings about (or is at least temporally contiguous with) such a reduction or state of 'relief,' so likewise is a reduction in the particular form of discomfort called anxiety effective in fixating behavior that is associated therewith.

Instead of first stating the systematic assumptions underlying the present study and then describing the experimental procedure and results, it seems preferable in this instance to

* The experimentation here reported was carried out under the writer's direction by Mr. Alex Mauro and Mr. Manuel Angulo. The writer likewise acknowledges his indebtedness to Mr. William D. Orbison, Miss Frances E. Campbell, and Mrs. Irene Anderson for other technical assistance.

[1] Interest has occasionally been expressed in the relation between various indices of 'emotion' and learning in particular situations. For example, Brown (**2**) has reported a correlation of .80 between the size of psychogalvanic response made by human subjects to serially presented nonsense syllables and the readiness with which these syllables were memorized. This finding, like many others that might be cited, is highly suggestive but has only an indirect bearing upon the problem at hand. For a discussion of the relation of emotion to 'education,' see Lund (**12**).

follow the opposite course. Since some of the problems here involved are relatively complex and as yet not completely solved, it is desirable to keep the theoretical part of this paper as concrete and unambiguous as possible. This end seems most likely to be attainable by going directly to an account of methodology and empirical findings, which can then serve as a fixed reference point for the subsequent theoretical analysis.

Apparatus and Subjects

The apparatus employed in the present study (see Fig. 1) is a modified version of the device first described by Culler, Finch, and Girden (4) for use with cats and later adapted by Hunter (11) for use with rats. It consists of a circular grill which is divided into eight individually electrifiable sections, enclosed by inner and outer walls of transparent celluloid (20 inches high). The grill itself is composed of small, radially placed, stainless steel rods and is five inches wide, with an internal diameter of 18 inches.

When a naive animal (rat or guinea pig) is first placed in this type of apparatus and current is applied to the particular section of the grill on which it happens to be located, it may churn its feet rapidly up and down, leap into the air, or engage in other equally futile agitated behavior. Very shortly, however, it is certain to try running (which sometimes occurs immediately); and since this behavior carries the animal off the charged section of the grill and thereby affords escape from shock, only two or three applications of the shock are required to make this reaction prompt and specific. Although running is not, therefore, a wholly 'unconditioned' response to shock under these circumstances, it can be so regarded for present purposes.

As a conditioned stimulus, a relatively pure tone (800 d.v. in pitch, 60 decibels above human threshold) was presented through headphones which were located three feet directly above the grill. All experimentation was carried out in a sound-proofed room, with visual and auditory cues from the experimenter excluded. The unconditioned stimulus (shock) was derived from a 60-cycle, 500-volt source of alternating current, with a standard limiting resistance in series with the subject (3,000,000 ohms for rats, 500,000 ohms for guinea pigs). A small rotary selector switch permitted prompt application of current to any of the eight sections comprising the grill.

Various conditioning set-ups of other types were tested before the one here described was adopted. It has an advantage over any arrangement demanding immobilization of subjects in that it saves time and avoids artifacts due to struggling. The modified activity cage described by Brogden and Culler (1) seems to be quite as satisfactory for use with guinea pigs as the circular grill arrangement, but it does not serve so adequately with rats for the reason that they have a tendency to keep running more or less continuously in it, which they do not do in the circular grill. The circular grill is likewise superior to the two-compartment device first described by Dunlap, Gentry, and Zeigler (6) and later modified by Warner (19), Hunter and Pennington (10), and others. In this latter type of apparatus strong conflict tends to be generated by the fact that the subject is forced to shuttle back and forth between only two positions, movement to either of which means going back to a place where punishment may have just been received. In the circular grill, on the other hand, an animal can always move on to a relatively 'new' position, one that is less strongly associated with punishment than the place it has just left and is therefore more acceptable. The only

complication that arose in connection with the use of the circular grill in the present investigation was that occasionally a rat tried to perch on one of the low crosspieces separating adjacent sections of the grill (see Fig. 1) and thereby avoid shock com-

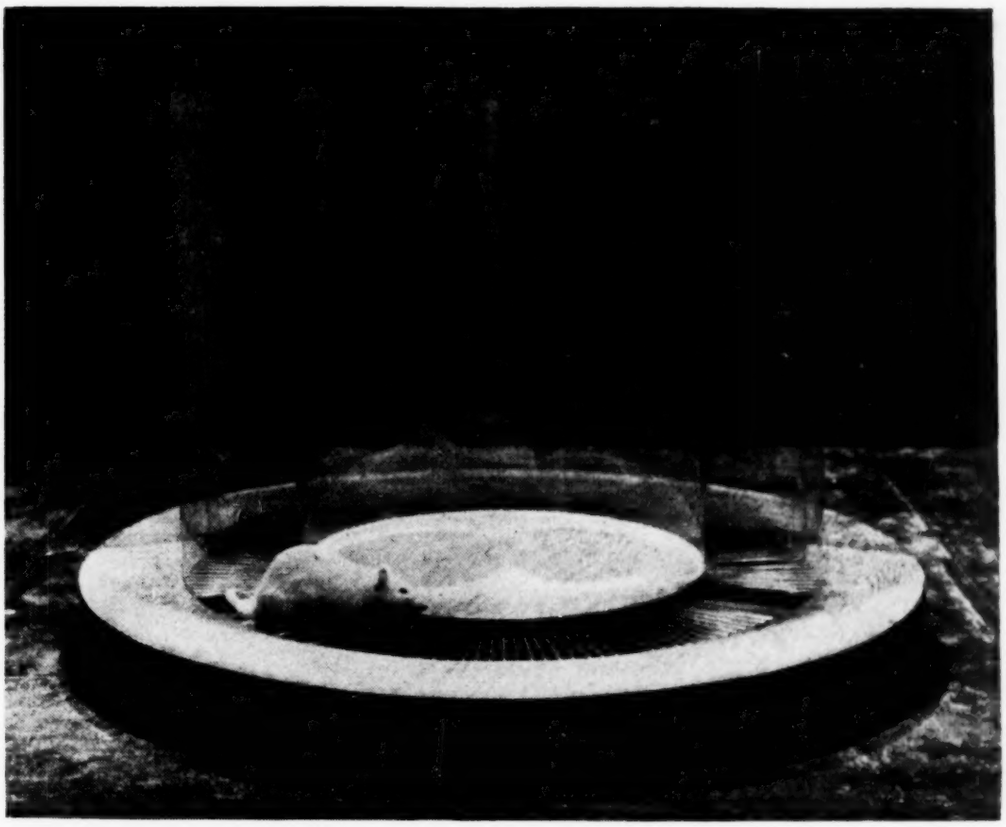


FIG. 1. Apparatus for the conditioning of a simple locomotor response.

pletely. When this happened, which was relatively infrequently, the rat was immediately dislodged by the experimenter.[2]

The 18 albino rats (12 males and 6 females) used in this investigation were obtained from the Wistar Institute (Experimental Colony Strain) and were approximately six months old. They had had no prior experimental use.

The 20 guinea pigs (10 males and 10 females) which were also employed were obtained from a local breeder. Little was known concerning them except that they were healthy young adults, with no prior laboratory experience.

PROCEDURE

Three distinct conditioning procedures were employed in this investigation and were carried out first with three groups of rats and then with three groups of guinea pigs.

The Group-I procedure was as follows. At the beginning of each experimental session, each animal was placed in the apparatus and allowed one minute of habituation. The conditioned stimulus, *i.e.*, tone, was then presented for five seconds, at the end of which time, if the rat had not moved out of the section in which it was sitting

[2] Following the completion of the study here reported, this defect was remedied by removing these separators and making the whole surface of the grill perfectly flat.

at the onset of the tone, shock was administered. On moving to another section of the grill, the rat escaped the shock and the tone was turned off by the experimenter. If, on the other hand, when the tone was presented, the rat made the required movement within the five-second interval following its onset, the tone was immediately terminated and no shock was applied.[3] Successive tones, or tone-shock combinations, were presented regularly every 60″. Training lasted for 10 days, each animal receiving 24 conditioning trials (tones or tone-shock combinations) each day.

The Group-II procedure was exactly the same as the Group-I procedure except for the fact that instead of successive tones, or tone-shock combinations, occurring at regular 60″ intervals, they were presented at three different intervals—15″, 60″, and 105″—in a revolving random order. During the course of each daily session of 24 trials, each of these intervals was used eight times, with the result that the *average* interval was exactly 60″. The average rate of presentation of stimuli was thus the same for Group-I subjects and Group-II subjects; but in the first case the rate of presentation was constant, whereas in the second case it was irregular.

The Group-III procedure was again exactly like the Group-I procedure save for one difference. In the Group-I procedure nothing happened to the subjects during the 60″ intervals between successive tones or tone-shock combinations. In the Group-III procedure, however, these intervals were punctuated by the interpolation of shocks every 15″ (*i.e.*, at 15″, 30″, and 45″ during each 60″ interval). These interpolated shocks were of the same intensity and were escapable in exactly the same way (*i.e.*, by flight to other sections of the grill) as were those which were paired with the tone. They could not, however, be avoided.

In short, the Group-II procedure differed from the Group-I procedure only in the matter of temporal regularity-irregularity of tone (or tone-shock) presentations, whereas the Group-III procedure differed from the Group-I procedure only in the matter of the filled-unfilled character of the intervals between successive presentations of tone (or tone and shock).

Results

The results, in terms of rate and amount of conditioning, obtained for the three groups of rats used in this investigation are represented graphically in Fig. 2. The ten successive points comprising each of the three curves show the average number of conditioned responses made by the six animals (4 males and 2 females) in each group during each of the 10 successive daily training sessions. It will be noted that the number of conditioned responses made by Group I is consistently superior to the number of conditioned responses mani-

[3] The practice of omitting shock when a conditioned response occurs, instead of having it invariably follow the conditioned stimulus, has been adopted because of the better conditioning obtained by Culler, Finch, Girden, and Brogden (5), Hunter (11), and others. However, the termination of the conditioned stimulus upon the occurrence of a conditioned response, instead of permitting it to continue for a fixed interval, seems not to have been previously described. Exploratory work indicates that this method may be superior to the conventional one; a more definitive study of this problem is now in progress.

fested by either Group II or Group III. The critical ratio of the difference between the amount of conditioning obtained in Groups I and III is 3.36.[4] This means, according to Fisher's (7) 'table of *t*', that the chances are more than 99 in 100 that

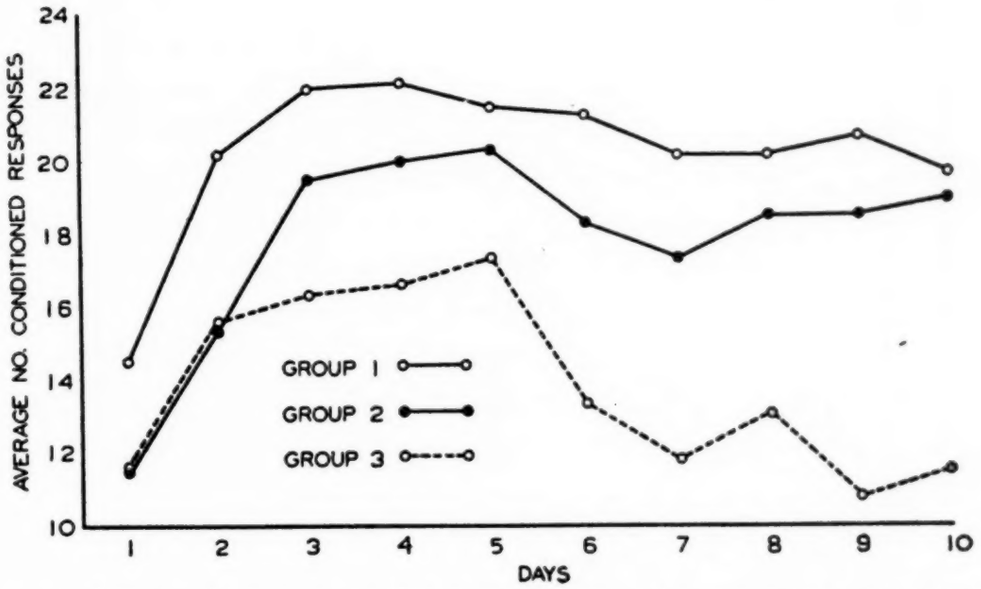


FIG. 2. Amount of locomotor conditioning obtained in three groups of rats in which reinforcement (anxiety-reduction) was systematically varied.

this difference is real (*i.e.*, caused by the conditions of the experiment). The critical ratio of the difference between the amount of conditioning obtained in Groups I and II is 1.63, which, in terms of '*t*,' means that there are about 85 chances in 100 that it is real. The critical ratio of the difference between Groups II and III is 1.88, which, again in terms of '*t*,' means that the chances are 90 in 100 that the difference is real. The average number of conditioned responses made by the Group-I animals during the entire 10-day training period was 200.0, by the Group-II animals 178.7, and by the Group-III animals 138.5.

The results obtained from the three groups of guinea pigs under the same conditions as employed with the three groups of rats are presented in Fig. 3. In general pattern, the find-

[4] This value was obtained by taking the *total* number of conditioned responses made by each animal during all 10 training sessions as the basis for computing the standard deviation of the distribution (σ). This practice of treating an animal's whole performance as a single 'test score' is followed throughout, except as otherwise indicated.

ings are very similar. In both types of subjects, the Group-I procedure gave the best conditioning, the Group-III procedure the worst, with the Group-II procedure intermediate. Although the rate of conditioning was much faster in the rats

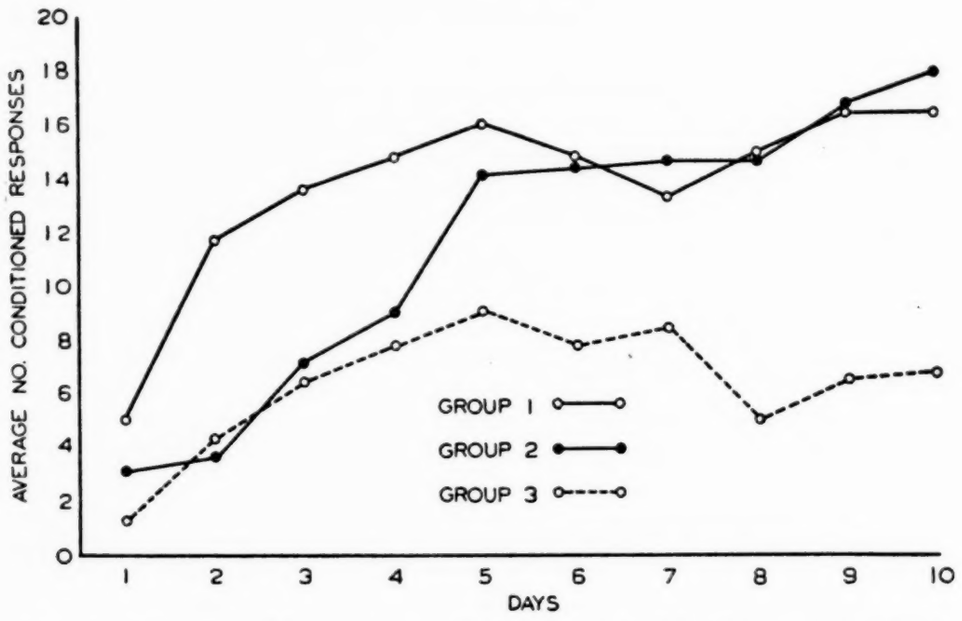


FIG. 3. Amount of locomotor conditioning obtained in three groups of guinea pigs in which reinforcement (anxiety-reduction) was systematically varied (as in Fig. 2).

than in the guinea pigs, the rats and guinea pigs alike showed a tendency for the conditioning in Groups I and II eventually to become equalized and for the conditioning in Group III to show an initial rise and then a marked deterioration. Statistical analysis indicates that in the case of the guinea pigs the chances are about 94 in 100 (CR. = 2.21) that the difference between Groups I and III is real and about 92 in 100 (CR. = 2.02) that the difference between Groups II and III is real. The difference between Groups I and II is, by the same criterion, much less reliable (CR. = .67), due to the virtual equalization of performance during the latter half of the training period. However, if one takes the performance of these two groups on the 2d, 3d, and 4th days of training, one finds that the chances are about 97 in 100 (CR. = 2.61) that the difference during this selected part of the procedure is real. The fact that this difference disappears as training

is prolonged does not mean that it is not real in the beginning. The average number of conditioned responses made by the Group-I animals during the entire training period was 137.1, by the Group-II animals 116.5, and by the Group-III animals 64.2.[5]

The relatively high '*t*' values and the marked similarity in the pattern of the differences obtained between the Group-I, Group-II, and Group-III procedures with both rats and guinea pigs as subjects seem unequivocally to support the conclusion that these differences are significant.

The activity that subjects manifest during the intervals between successive paired presentations of conditioned and unconditioned stimuli has been regarded by most investigators as a nuisance, something to be either ignored or designedly minimized by whatever means available. In the present study this activity has been taken into account and carefully recorded, on the assumption that it is related to and may throw light upon the dynamics of the conditioning process as a whole. Figure 4 and Fig. 5 show for rats and guinea pigs, respectively, the average number of sections of the circular grill traversed per daily training session by the animals in Groups I, II, and III.[6] Although differing considerably in absolute value, the findings for the two types of subjects are here again very much alike. In both cases, the Group-III animals consistently show the greatest amount of activity, the Group-I animals show the next greatest amount, and the Group-II animals show the least. The critical ratios of the differences between these groups are 3.37 (I–II), 4.41 (II–III), and 1.78 (I–III) for the rats and 3.38 (I–II), 9.30 (II–III), and 6.88 (I–III) for the guinea pigs. There can thus be little question of the reliability of these differences.

[5] It will be noted that these figures are considerably lower than those obtained from the corresponding groups of rats. The superiority of the performance of the latter type of subjects is further indicated by their lower variability. The rats had coefficients of variability of 6.1 (Group I), 14.9 (Group II), and 21.8 (Group III), as against 44.3 (Group I), 36.4 (Group II), and 73.3 (Group III) for the guinea pigs.

[6] These averages are based upon all the activity manifested by each animal during the experimental session, including not only the 'spontaneous' locomotion that occurred between trials but also the responses resulting directly from application of the conditioned and unconditioned stimuli.

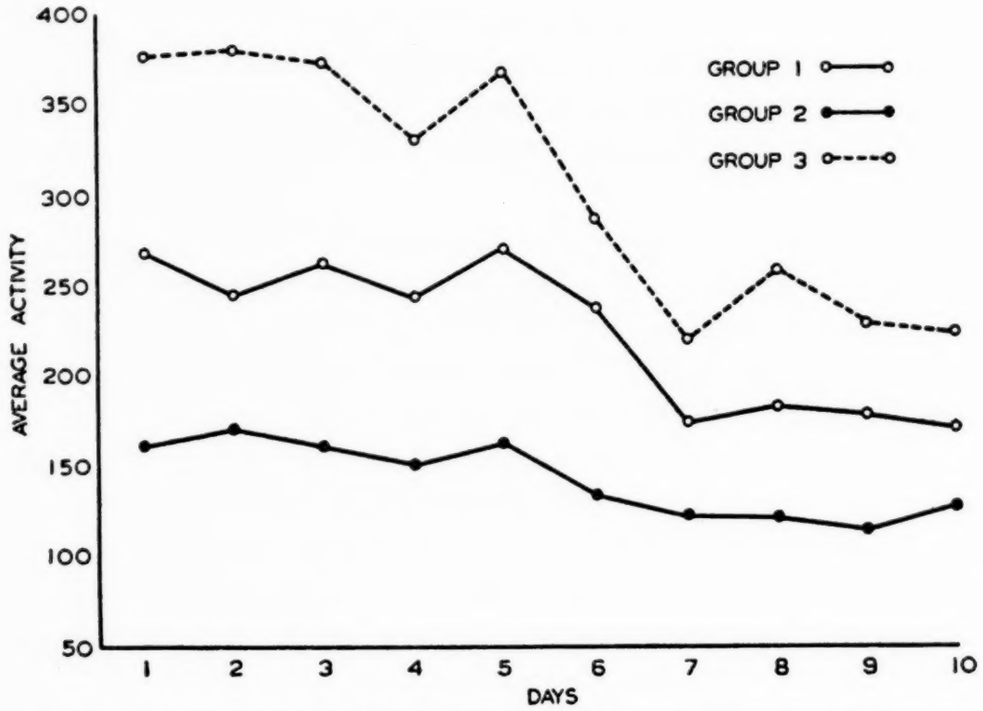


FIG. 4. Average amount of total locomotor activity manifested per daily training session by rats during the development of a conditioned locomotor response (compare Fig. 2).

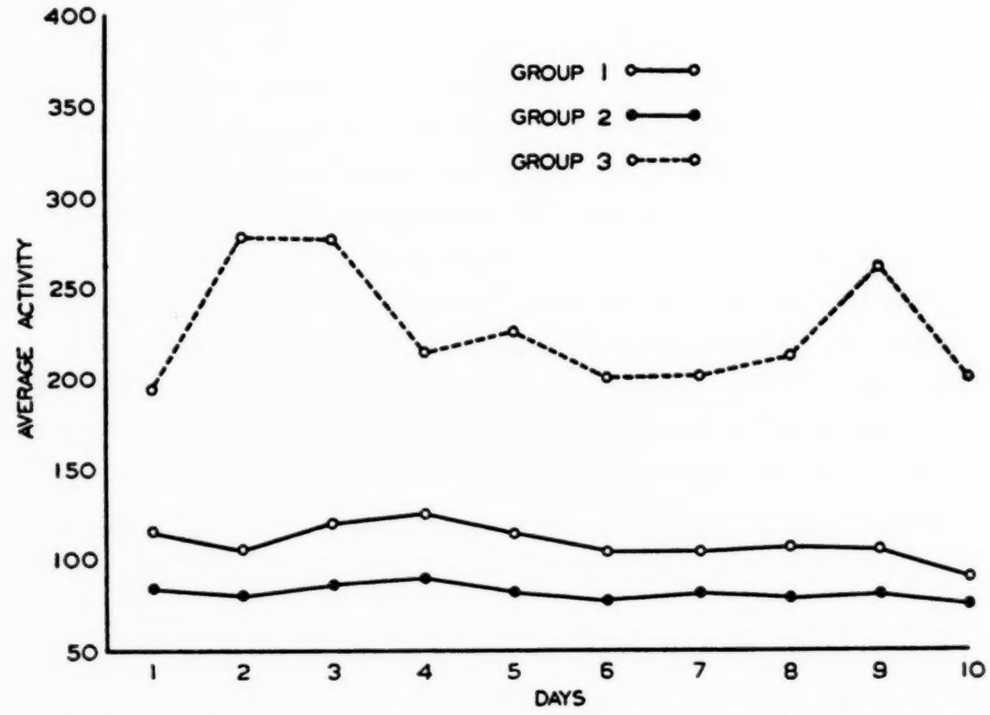


FIG. 5. Average amount of total locomotor activity manifested per daily training session by guinea pigs during the development of a conditioned locomotor response (compare Fig. 3).

With the hope of securing an additional basis on which to compare the amount of conditioning obtained in both rats and guinea pigs by the three procedures here employed, all subjects were submitted, on the day following the completion of the 10-day training period, to a standard 24-trial extinction series. Each animal was put into the apparatus, habituated as usual, and then presented with 24 tones at regular one-minute intervals. If, within five seconds after the onset of tone, an animal moved to another section of the grill, the tone was promptly turned off and a conditioned response recorded; if the animal did not move, the tone was left on for the full five-second period and then turned off, with no shock in either case. Figure 6 shows the extinction curves for the three

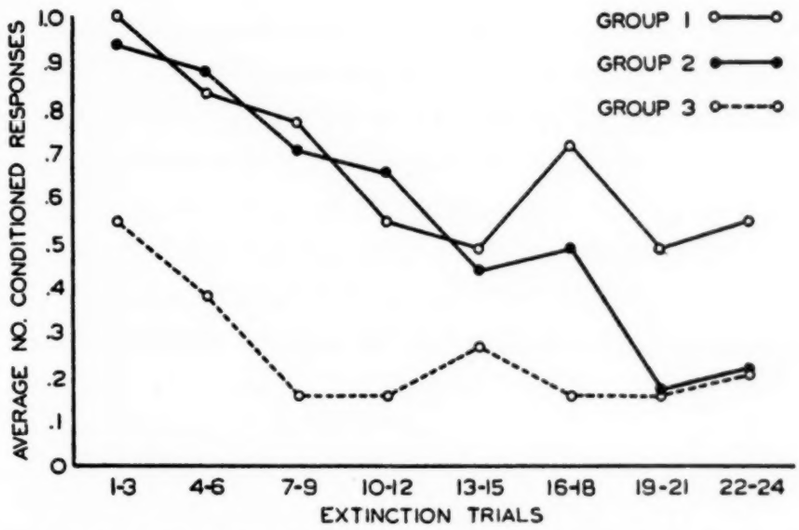


FIG. 6. Extinction of a conditioned locomotor response in rats.

groups of rats, and Fig. 7 shows the corresponding curves for the three groups of guinea pigs. In general, they tell much the same story as do the acquisition curves (Fig. 2 and Fig. 3). Resistance to extinction is markedly greater in both the Group-I and Group-II animals than in the Group-III animals. In the case of the rats, the resistance of the Group-I animals seems to be slightly greater than that of the Group-II animals; but in the case of the guinea pigs this relationship is reversed. No attempt has been made to determine the statistical reliability of these differences.

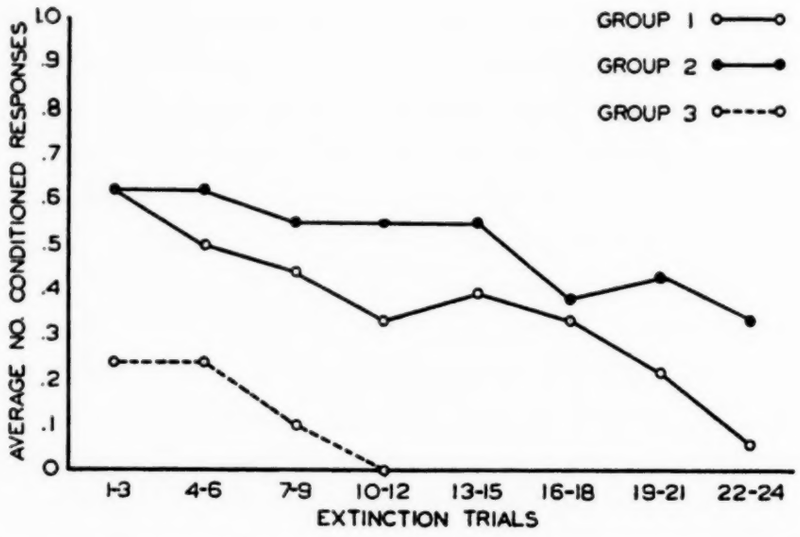


FIG. 7. Extinction of a conditioned locomotor response in guinea pigs.

The total locomotor activity manifested during the extinction trials by both rats and guinea pigs is shown, respectively, in Fig. 8 and Fig. 9. Here some marked dissimilarities are evident. Although the Group-II animals hold the intermediate position in both cases, the Group-I and Group-II curves are exactly reversed. The Group-III rats continue to show the greatest activity during extinction, as they did during acquisition; but the Group-III guinea pigs show the least

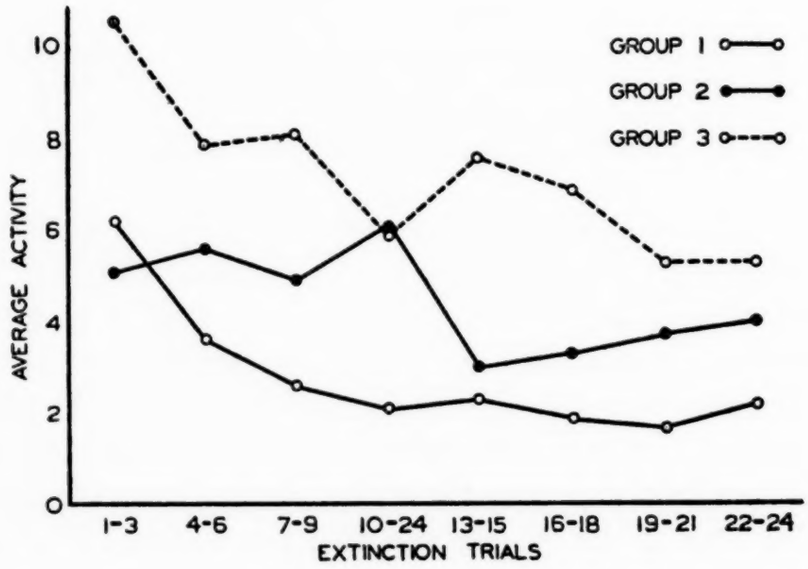


FIG. 8. Average amount of total locomotor activity manifested per trial during extinction of a conditioned locomotor response in rats (compare Fig. 7).

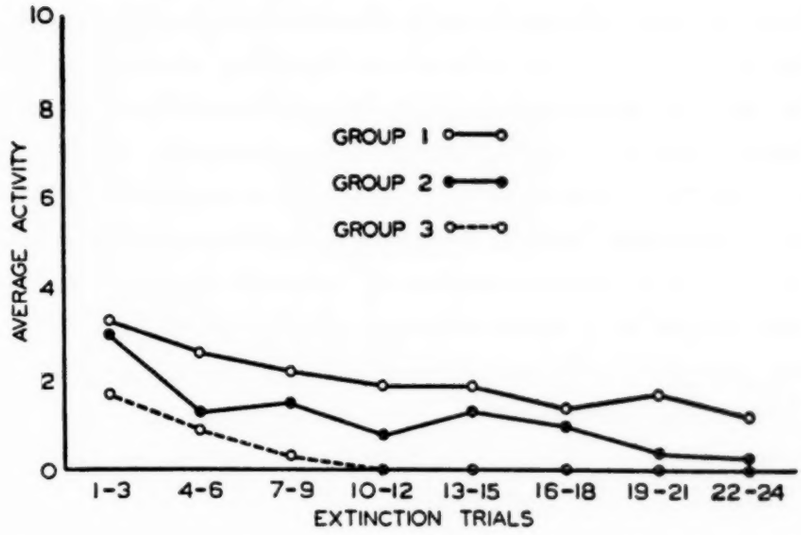


FIG. 9. Average amount of total locomotor activity manifested per trial during extinction of a conditioned locomotor response in guinea pigs (compare Fig. 8).

activity during extinction, instead of most as they did during acquisition. Still other differences are to be noted, which will be discussed later.

DISCUSSION

Evidence which has been presented elsewhere (**15**) indicates that when a stimulus is recurrently presented to an organism at regular temporal intervals, expectation of (preparedness for) this stimulus mounts as the time approaches for its presentation, suddenly drops when the expected stimulus occurs (and is reacted to), then begins to mount again, and so on. When this phenomenon of expectancy, or preparatory set, is sufficiently highly developed to be 'uncomfortable,' *i.e.*, takes on motivational properties in its own right, its reduction should, on the basis of the hypothesis stated at the outset of this paper, provide a 'satisfying state of affairs' and thereby 'reinforce' accompanying behavior. As a corollary of this proposition, it is to be supposed that, within limits, the greater the extent of such a drop in preparatory tension (dread, anxiety) at any given point in time, the greater its reinforcing, or learning-producing, potentialities.

Again on the basis of evidence reported in the study cited in the preceding paragraph, it is clear that if a stimulus is presented at irregular intervals, the drop in preparatory ten-

sion that occurs immediately thereafter is less than if the stimulus occurs at a standard, unvarying rate. It is consequently to be inferred from this observation that if the hypothesis under investigation here—namely, that reduction in preparatory tension, or anxiety, is a reinforcing state of affairs—is correct, better learning should occur in a situation involving regular presentation of stimuli than in a situation involving irregular presentation. The superiority of the conditioning manifested by the Group-I animals in the present study (regular presentation of stimuli) over the conditioning manifested by the Group-II animals (irregular presentation of stimuli) confirms this inference.[7]

Although these results are consistent with the stated hypothesis, the possibility immediately suggests itself of explaining them in another way. Pavlov (**17**) and other investigators have shown that conditioned responses can be established on the basis of the mere passage of time. These so-called 'trace' conditioned reactions are, as one would intuitively suppose, more readily established to a temporal interval that is constant than to one that is irregular. On the basis of these observations it might be maintained, therefore, that the Group-I animals in the present investigation responded more frequently to the presentation of tone alone than did the Group-II animals solely for the reason that, on the basis of so-called temporal conditioning, the response in question (running from one section of the grill to another) was more nearly ready to occur as a 'trace' conditioned response at the time the tone was presented in the one group than in the other. In other words, with regular temporal intervals between successive stimulations, the response in question would be on the verge of occurring anyway, as a 'trace' conditioned response, and would therefore be relatively easy to 'trip' by the tone; whereas, with irregular intervals between successive stimula-

[7] Preliminary experimentation has indicated that this difference appears only if the absolute duration of both the regular and the irregular intervals is relatively brief. If, for example, an interval of 120 seconds (instead of 60 seconds) were permitted between trials in the regular order of presentation and if intervals of 75, 120, and 165 seconds (instead of 15, 60, and 105) were permitted between trials in the irregular procedure, any difference in results would probably be much reduced if not obliterated entirely.

tions, the response would be less highly prepared at any particular point in time at which the tone happened to occur, with the result that the tone would be less effective in eliciting the response. If this simple view of the matter were correct, there would be no need to posit differential anxiety-reduction to account for the differing amounts of conditioning obtained in the Group-I and Group-II animals.

The procedure employed with the Group-III animals was specifically designed to test the soundness of the foregoing explanation. It will be recalled that in the case of the Group-III subjects, tone (or the tone-shock combination) occurred, as in the Group-I procedure, at regular 60″ intervals but that shocks were interpolated every 15″. The Group-III animals were thus forced to make a response, *i.e.*, move to another section of the grill, regularly every 15″, one fourth of the time in response to tone (or tone and shock) and three fourths of the time in response to shock alone. In view of both the brevity and the regularity of this 15″ interval, temporal conditioning should, theoretically, have occurred especially quickly; and if the development of conditioned reactions to tone were primarily dependent upon this underlying temporal conditioning, then the Group-III animals should have shown superior learning. However, this was not the case: the conditioning obtained in the Group-III animals was decidedly inferior.[8]

[8] After this paper was in press, Mr. William D. Orbison called the writer's attention to a possible alternative explanation of this finding. Since the interpolated shocks could not be circumvented by anything the subjects could do, whatever they did just preceding the shocks was necessarily "punished." However, as Brogden, Lipman, and Culler have pointed out (The rôle of incentive in conditioning. *Amer. J. Psychol.*, 1938, **51**, 109–117), a shock administered to a small animal such as the rat or guinea pig from a grill is probably less "punishing" if the animal assumes a preparatory crouch than if the animal is running at the moment of shock onset. Thus, a recurrent shock which is inescapable tends to be met by immobility; whereas, a shock which can be averted by anticipatory running tends to perpetuate this type of behavior. In the procedure employed with the Group-III animals in the present study, conflict may, therefore, have been generated as follows: whereas the interpolated (inescapable) shocks tended to immobilize the animals, the shock which followed the tone only if the animal did not make a conditioned response tended to produce anticipatory running. A glance at the Group-III curves in Fig. 2 and Fig. 3 suggests that during the early part of the 10-day experimental period the subjects were trying to make an adjustment to the situation by means of the latter type of

These results do not, of course, deny the reality of temporal conditioning, but they do indicate that it alone cannot provide a comprehensive explanation of the results obtained in the present study. On the other hand, the anxiety-reduction hypothesis does provide such an explanation. Speaking loosely, it may be said that the poorness of the learning obtained in the Group-II and Group-III animals was due to the fact that they were given less opportunity to 'let down' between trials than were the Group-I animals. In other words, they were kept in a more or less chronic state of apprehension or suspense. In the Group-II animals this was accomplished by 'keeping them guessing' as to when each succeeding stimulation was to occur. In the Group-III animals, this element of uncertainty was presumably absent, but the stimuli came in such rapid succession that one stimulus had no more than occurred when the next one began to be 'expected.'[9] In either case, conditioned reactions to the tone were relatively 'unrewarding' in that little was gained in the way of anxiety-reduction when they occurred; whereas, in the case of the Group-I animals each response was followed by a relatively long, unvarying period of no stimulation, with correspondingly great anxiety-reduction.

The fact that the Group-II animals showed conditioning which, although inferior to that of the Group-I animals, was markedly superior to that of the Group-III animals indicates,

---

behavior; but since this did not prove satisfactory, they soon resorted to the other mode of adaptation, namely, anticipatory crouching, with a resulting deterioration in their "conditioned-response" scores. The fact that the activity record of the Group-III rats broke sharply (see Fig. 4) on the same day that their conditioned-response score dropped most marked, namely, on the sixth day, would tend to support this interpretation, although the activity record of the Group-III guinea pigs showed no such slump (see Fig. 5). Although this "conflict" explanation of the inferior conditioning obtained with the Group-III animals is somewhat different from the one proposed in the text, it is similar in that it, too, involves a law-of-effect view of the conditioning process instead of the traditional associationistic theory. Furthermore, there appears to be such an intimate connection between anxiety and conflict that the two alternative hypotheses here proposed may ultimately reduce to essentially the same thing.

[9] That the Group-III procedure kept apprehension chronically high is indicated by the fact that the Group-III rats made repeated efforts to 'leave the field' by trying to jump out of the apparatus. This behavior was not quantified but was very striking.

presumably, that the Group-II procedure was such as to permit a degree of anxiety-reduction between trials that was less than that permitted by the Group-I procedure but greater than the permitted by the Group-III procedure. Direct measurement of the respective amounts of anxiety-reduction permitted under the conditions of the three procedures would have been a decided advantage, but this did not seem possible. Superficially it might appear that the total amount of activity displayed by the various groups of animals (Fig. 4 and Fig. 5) would provide an empirical index of the amount of tension they were under. It is true that the Group-III animals, which were conjecturally kept most anxious, showed a uniformly higher level of total activity than did the other two groups. However, it is impossible to say to what extent this greater activity was due to anxiety and to what extent it represented responses made directly to the numerous interpolated shocks. More meaningful perhaps is the fact that the Group-I animals showed a consistently higher level of total activity than did the Group-II animals: Two alternative interpretations of this difference are possible. It can be assumed, on the one hand, that greater activity represents greater anxiety and that the Group-I animals were consequently more anxious; or it can be assumed that greater activity represents a relative absence of anxiety, in which event the Group-I animals would be viewed as less anxious. Observations made on rats and guinea pigs in other experimentally produced danger situations suggest that the latter supposition may be the more valid one. However, additional experimentation, aimed specifically at this problem, will be necessary before a completely reliable decision can be reached on this score.

Another objection that may be raised to the interpretation of the findings of the present study in terms of differential anxiety-reduction is that anxiety itself is a learned phenomenon and should not therefore be used as a basis for explaining further learning. The writer (**16**) is in full agreement with the proposition that anxiety is a learned phenomenon, but it does not follow from this that it cannot function as a motive and its reduction as a reinforcing agent for further learning.

Numerous recent studies in the field of motivation have made it clear that in man and the lower animals alike, there are both innate and derived forms of motivation. The position here taken is that anxiety falls into the latter category. Just how the cyclic rise and fall of anxiety, or preparatory tension, posited above as a condition of recurrent noxious stimulation is acquired is an important problem, but its solution does not necessarily have to be completed before use can be made of the phenomenon itself as an explanation of other findings.

It is unfortunate that the hypothesis here advanced lacks the great simplicity of the associationistic theory of conditioning advocated by Pavlov and his followers. There is, however, no point in venerating a theory because of its simplicity if it is demonstrably inadequate to account for all the relevant facts. Some of the many weaknesses of the view that mere paired presentation of stimuli provides the necessary conditions for conditioned-response learning have been previously enumerated (**14**) and need not be repeated here. It is, however, instructive to note that under the conditions of the present experiment, best conditioning occurred precisely in those animals which received *fewest* paired presentations of tone and shock. Since the results presented in Fig. 2 and Fig. 3 are in terms of the average number of conditioned reactions shown by the various groups of animals during the course of each daily training session, and since a conditioned response is defined as one that occurs to a tone which is not followed by shock, it is axiomatic that the animals showing the best conditioning must have received the smallest number of paired presentations of tone and shock. In view of this fact it follows that the goodness of learning in a situation of this kind is not determined exclusively or even predominantly by the temporal pairing of stimuli but by some other factor. The assumption here made is that this other factor is the presence or absence of a reinforcing state of affairs as defined by the law of effect. The temporal contiguity of stimuli seems important only from the point of view of determining what stimuli will become functionally connected with what responses; it presumably does not determine whether such connections will or will not occur.

Two other implications are to be briefly noted. It is well established that human beings memorize verbal material more efficiently with distributed practice than with massed practice. Without attempting to evaluate or compare the virtues of other theories which have been advanced to account for this phenomenon, it is worth observing that a possible explanation follows fairly directly from the foregoing discussion. In the learning of verbal material, motivation is ordinarily supplied, not by primary motives such as hunger or thirst, but by some form of tension or mild anxiety (anticipation of being ridiculed for poor performance, being called stupid, or the like). Reinforcement may be assumed to occur when the subject reproduces a correct item and is permitted to know that it is correct. Let this state of affairs be thought of as involving tension-reduction. However, each successive task that is set the subject involves the possibility of failure, *i.e.*, arouses new tension, and the more rapid the rate at which these tasks are presented, the less the tension will drop after each correct response. Parents soon learn the wisdom of spacing their demands upon small children for acquiring new responses. Each demand for learning produces an increased tension (renewed anticipation of failure, punishment, etc.), and if such demands come in too rapid succession, tension will remain chronically so high that the proper tension-reduction ('success') cannot be experienced after a good performance to insure its fixation. Laboratory evidence in support of this theory as to why distributed practice is so much superior to massed practice in certain types of learning situations will be discussed in a later report.

The other implication of the present findings that seems noteworthy bears upon some experimental results reported by Calvin (**3**). This investigator found that human subjects show better eyelid conditioning when an irregular temporal interval is allowed between successive paired presentations of the conditioned and unconditioned stimulus than when a regular interval is employed. In a previous publication (**15**), the present writer has suggested that this finding may be due to the fact that human beings tend to develop a 'negative

attitude' in most conditioning situations and that the more precisely they can judge when a conditioned stimulus is going to occur, the more effectively they can guard against the occurrence of a conditioned response. In the same study it was also conjectured that since the complex social factors that are presumably responsible for the development of negative attitudes on the part of human subjects in conditioning situations are not operative in animals, the latter should show better rather than poorer conditioning when the rate of stimulus presentation is regular. The results of the present investigation support this hypothesis.

Before concluding this paper it is perhaps worth calling attention to the fact that the curves for conditioning (Fig. 2 and Fig. 3) and those for total activity follow divergent courses. Whereas the former present a general pattern of increment, the latter show more or less decrement. The conditioning process is thus to be seen as one of progressive differentiation and specification of response. In describing the development of a conditioned paw-retraction response in dogs, Culler, Finch, Girden, and Brogden (**5**) say:

> During the early periods of training, the shock often evokes vigorous and varied activity (yelping, jerking, twisting, evacuation, to mention only the more evident manifestations of fright, escape or defense). . . . In these earliest stages we find no trace of a localized reaction (foot-withdrawal); the animal's activity all belongs to the general pattern of escape-defense. Now and then it misses the shock, but only by accident; the right forepaw moves in accord with general somatic activity and not in any adaptive manner. . . . Later, on the contrary, when the animal has achieved an adaptive response (foot-withdrawal prior to shock) there is but little resemblance. The localized, individuated response of a well-trained dog is as unlike his original excited, diffuse behavior as two acts can well be (p. 224).

Although it is evident that the process of individuation was to some extent involved in the conditioning of the locomotor response investigated in the present study, it was much less pronounced than in the situation just described. This fact probably accounts in large measure for the relatively greater ease of conditioning of a gross locomotor response than of a more limited and specific reaction such as paw-retraction. Schlosberg (**18**) experienced great difficulty in obtaining a stable conditioned leg-withdrawal response in rats, and the

present writer has found in unpublished work that it is correspondingly hard to condition a type of pedal-pressing response which has been described elsewhere (**13**). If a generalization is here warranted it would seem to be that the more a conditioned response is permitted to resemble the response initially made to the unconditioned stimulus, the more rapidly conditioning proceeds. There seems also to be a rather nice correlation between the phylogenetic position of an organism and the extent to which conditioned responses can be made to differ from their unconditioned prototypes. Man undoubtedly stands at the top of this scale. Whereas the conditioned responses which lower animals learn to make to signals of danger are, for example, usually not very different from those which they would make if the signified trauma were actually present, man, on the other hand, is readily capable of acquiring responses to danger signals that differ drastically from the behavior which would be manifested in the face of the danger itself. This difference is, however, apparently one of degree and is in no sense absolute.

Summary

The primary purpose of the present study was to test the hypothesis that anxiety-reduction acts as a reinforcing state of affairs (as defined by the law of effect). To this end three experimental procedures, designed to provide three different degrees of anxiety-reduction, were carried out, with both rats and guinea pigs as subjects. The results indicate that anxiety-reduction is positively correlated with learning, thereby supporting the supposition that it is a reinforcing agent.

Various implications of this finding are discussed and contrasted with alternative interpretations suggested by the conditioned-reflex (associationistic) conception of the learning process. The latter is found deficient because of its emphasis upon the importance of mere temporal contiguity (paired presentation) of stimuli and its neglect of the factors of motivation and motivation-reduction.

(Manuscript received May 28, 1940)

References

1. Brogden, W. J., and Culler, E., Device for the motor conditioning of small animals, *Science*, 1936, **83**, 269–270.
2. Brown, C. H., The relation of magnitude of galvanic skin responses and resistance levels to the rate of learning, *J. Exp. Psychol.*, 1937, **20**, 262–278.
3. Calvin, J. S., *Decremental factors in conditioned-response learning* (Yale Ph.D. Dissertation, 1939).
4. Culler, E., Finch, G., and Girden, E., Apparatus for motor conditioning in cats, *Science*, 1934, **79**, 525–526.
5. Culler, E., Finch, G., Girden, E., and Brogden, W., Measurements of acuity by the conditioned-response technique, *J. Gen. Psychol.*, 1935, **12**, 223–227.
6. Dunlap, K., Gentry, E., and Zeigler, T. W., The behavior of white rats under food and electric shock stimulation, *J. Comp. Psychol.*, 1931, **12**, 371–378.
7. Fisher, R. A., *Statistical methods for research workers*, London: Oliver and Boyd, 1938, pp. 356.
8. Freud, A., *The ego and the mechanisms of defence*, London: Hogarth Press, 1937, pp. 196.
9. Freud, S., *The problem of anxiety*, New York: Norton and Co., 1936, pp. 165.
10. Hunter, W. A., and Pennington, L. A., A new apparatus and method for training the rat in auditory discrimination problems, *Science*, 1939, **89**, 87–88.
11. Hunter, W. S., Conditioning and maze learning in the rat, *J. Comp. Psychol.*, 1935, **19**, 417–424.
12. Lund, F. H., *Emotions—their psychological, physiological, and educative implications*, New York: Ronald Press, 1939, pp. 305.
13. Mowrer, O. H., An experimental analogue of 'regression' with incidental observations on 'reaction-formation,' *J. Abn. (Soc.) Psychol.*, 1940, **35**, 56–87.
14. Mowrer, O. H., Preparatory set (expectancy)—a determinant in motivation and learning, *Psychol. Rev.*, 1938, **45**, 62–91.
15. Mowrer, O. H., Preparatory set (expectancy)—some methods of measurement, *Psychol. Monogr.*, 1940, **51**, pp. 53.
16. Mowrer, O. H., A stimulus-response analysis of anxiety and its role as a reinforcing agent, *Psychol. Rev.*, 1939, **46**, 553–565.
17. Pavlov, I. P., *Conditioned reflexes*, London: Oxford University Press, 1927, pp. 430.
18. Schlosberg, H., Conditioned responses in the white rat, *J. Genet. Psychol.*, 1934, **45**, 303–335.
19. Warner, L. H., The association span of the white rat, *J. Genet. Psychol.*, 1932, **14**, 57–90.

AN ATTEMPT TO CONDITION THE PUPILLARY RESPONSE

BY C. H. WEDELL, F. V. TAYLOR AND A. SKOLNICK

Princeton University

This paper describes an attempt to condition the pupillary response to an auditory stimulus. Conditioning of the pupil is not new. It was reported as early as 1916 by Watson (**10**). In spite of the many experiments which have been performed since then (**1**, **2**, **4**, **6**, **7**, **8**, **9**), there is no clear understanding of the necessary and sufficient conditions for the production of this conditioned response. In fact, there is considerable disagreement as to whether the pupil can be conditioned at all (**8**, **9**). In view of the contradictory findings in this field, the recently published experimental results of Baker (**1**) and Metzner and Baker (**7**) are startling in their implications. Not only do they report conditioning of the pupillary response but also they claim that conditioning was established in two trials provided the auditory stimulus was below the 'conscious limen' (**1**, p. 13). Baker presents evidence to show that if the auditory stimulus is below this limen, a conditioned response of the 'disturbance' type can be shown to exist at the second simultaneous presentation of light and tone, and the 'final form' response appears at the third trial. This final form response is very stable and very specific. He also states that if the auditory stimulus is supraliminal, a much longer series of paired presentations (from 29 to 105 trials) is necessary to obtain the CR.

These results seemed important enough for the theory of conditioning to warrant a repetition of the experiment. The results of this repetition are reported in the following pages.

Apparatus

The apparatus (Fig. 1) consisted of three parts: the pupillometer, the recording device, and the stimulating instruments. The pupillometer was a modification of similar devices described by Baker (1), and Ferree and Rand (3), and utilized the

principle of the transparent mirror. A telescope was directed upon the eye of the subject. The objective of the telescope was 50 cm from the subject's eye and the telescope had a magnification of about 6 diameters at this distance. Between the subject's eye and the lens of the telescope was interposed a transparent piece of glass at an angle of 45 degrees to the axis of the scope. This glass reflected the image of two black threads. These threads thus appeared to be superimposed upon the pupil of the

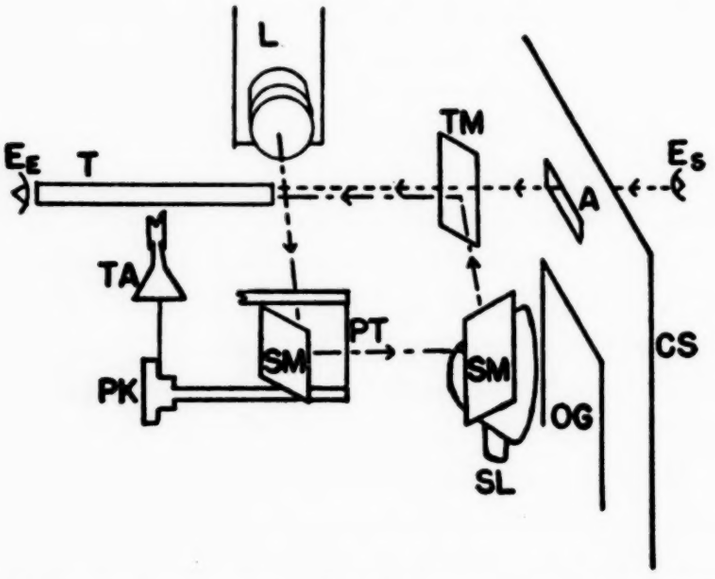


FIG. 1. Diagram of apparatus. The relative position of the various parts is the same as in Fig. 2. Light from a source (L) is reflected from the silvered mirror (SM) and throws a shadow of the pupillometer threads (PT) on a second silvered mirror (SM). The image of the threads travels from this mirror to the transparent mirror (TM) and is directed through the telescope (T) into the eye of the experimenter (Ee). At the same time, light from the 60 watt bulb in a reflector (SL) is diffused through the opal glass (OG) in the cardboard screen (CS) and is reflected from the subject's eye (Es) through the opening in the screen (A), the transparent mirror (TM) and the telescope (T) to the experimenter's eye (Ee). The distance between the pupillometer threads can be varied by turning the pupillometer knob (PK) which in turn actuates the tambour (TA).

subject. The distance between them could be varied by the experimenter by turning the knob of the pupillometer so that variations in the size of the pupil could be followed easily. Turning the knob of the pupillometer also actuated an air tambour, and the variation of air pressure was conducted to another tambour which controlled the recording pen on the kymograph located in another room. The kymograph was a continuous paper Ohio kymograph on which was recorded with ink-writing pens the onset of the two stimuli, light and sound; the time in seconds; and the variations in pupil size as transmitted from the pupillometer.

The light source for the unconditioned stimulus was a 60-watt frosted glass bulb located behind an opal glass screen as described below. A change in intensity from 3.5 to 20 foot candles by means of a General Radio Variac transformer was the unconditioned stimulus. The intensities of the UCS were the same as those used by Baker. Calibration of the light intensity was made with a Weston exposure meter which was

held in the position of the eye of the subject. It is realized that this measurement was probably not very accurate but it is believed to be accurate enough for the purpose of the experiment.

Two auditory stimuli were used as conditioning stimuli. With some subjects a tone of 1024 cycles was employed, and with others the noise from a buzzer. The tone

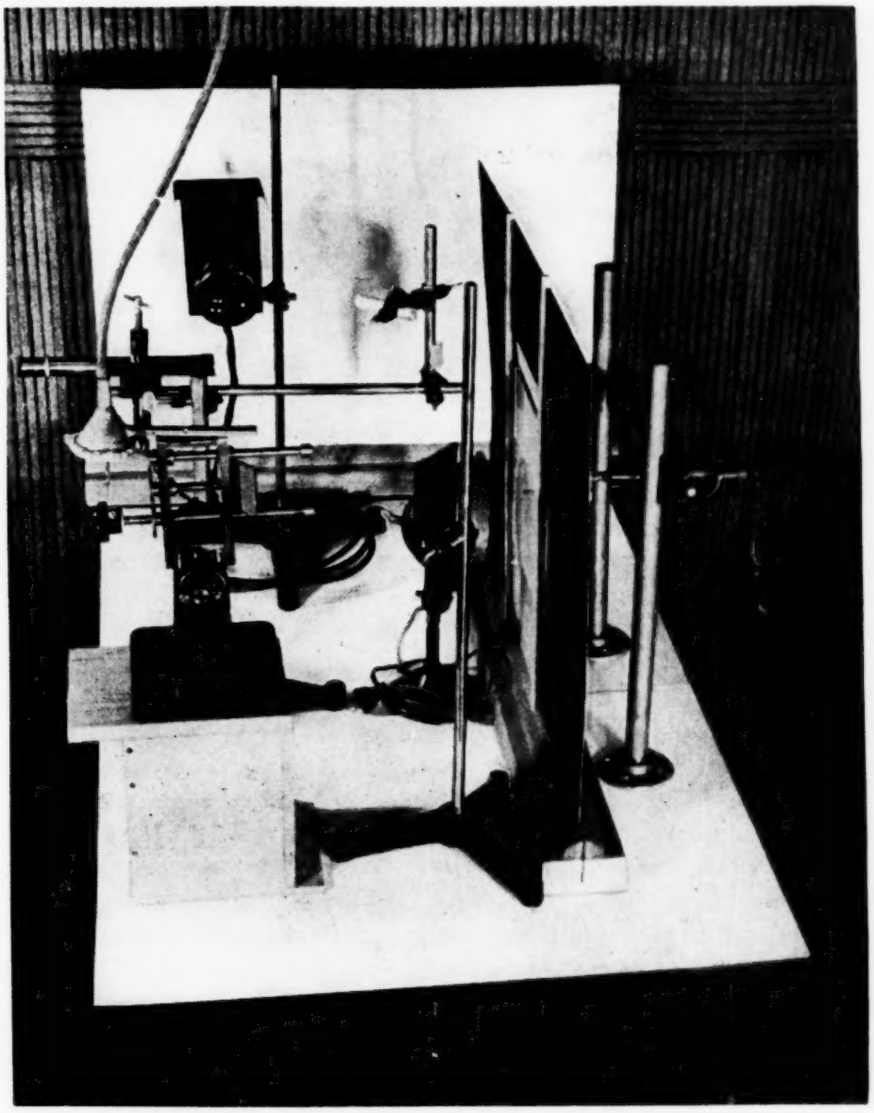


FIG. 2. Photograph of apparatus. See Fig. 1 for interpretation.

was generated by a Western Electric 2A audiometer and passed through an attenuator which made it possible to control the intensity in steps of one decibel. The buzzer was a small commercial instrument, emitting a noise of predominantly low frequency, whose intensity was controlled by the same attenuator. Whichever stimulus was used, it was delivered to the subject through a headphone. The tone was used with the first seven subjects. Later, it was learned that the Seashore audiometer which Baker employed delivers a noise rather than a tone; therefore a buzzer was used instead of the 1024 cycle tone with the last four subjects.

The subject and one of the two experimenters sat at opposite sides of a table in a

soundproof room. Directly in front of the subject was a large, black cardboard screen. Covering an opening in the center of the lower half of the screen was a piece of opal glass, 22 cm by 32 cm in size, behind which was the 60-watt frosted glass light bulb which served as the source of the unconditioned stimulus. Just above the opal glass was a small opening at the height of the subject's eyes through which he could see the objective of the telescope and a small piece of white cardboard with a black dot in the center of it which served as the fixation point. The fixation point was 19 cm from the eye of the subject. This is a rather short fixation distance but was used because of the desirability of an exact duplication of Baker's conditions. The subject's head was immobilized by having him grip a small dowel-stick with the teeth. Because of the small opening in the screen and the distribution of light, the experimenter and the pupillometer on the other side of the table were only vaguely and incompletely visible to the subject.

A second experimenter operated the stimulating and recording apparatus in an adjacent room and controlled the time of onset of all stimuli. He also watched the kymograph and made all necessary records. Communication between the two experimenters was by telephone.

There was no instrumental noise of any kind in the soundproof room and the experimenter operating the pupillometer was ignorant of the time of onset of the conditioned stimulus. However, in a few trials toward the latter part of the experiment a signal was given to the experimenter operating the pupillometer so that he could pay particular attention to the behavior of the pupil during the critical interval following the presentation of the CS. This precaution was taken to make certain that a very slight CR would not go unnoticed.

Magnification.—Because of the fact that the movement of the pupillometer knob was transmitted to the recording pen on the kymograph by tambours actuating a column of air and the fact that the recording pen swung in an arc, the magnification of the change in the size of the pupil varied in amount. At the beginning of the swing of the recording pen, 1 mm increase in diameter of the pupil caused a movement of 3.5 mm of the recording pen. Midway in the swing of the pen, a 1 mm change in pupil size caused the writing point to move 10 mm, and at the other end of the swing the movement of the pen again decreased to 3 mm per 1 mm change in pupillary diameter. These magnifications are not large, but they are large enough so that the variations in the size of the pupil not associated with any observable stimulus-change were easily distinguished from the changes caused by the unconditioned stimulus.

Method

Each subject's limen for the auditory stimulus to be employed was obtained before the first session with that subject. It was desired to use a method of threshold determination which would result in a very low limen. After some exploration, it was found that the following method gave a limen that was lower by five or ten decibels than that resulting from the usual method of limits. The method consisted of first obtaining the threshold roughly by the method of limits. Then, beginning about five decibels above this value, the experimenter set the tone-switch either on or off and signaled the subject. The subject listened and reported whether the tone was on or off. This procedure was repeated at intensities diminishing in steps of 1 db until the limen was reached. The limen was defined as that point above which the responses were better than chance, *i.e.* better than 5 correct out of 10 trials.

The stimulation times were the same as those used by Baker. The conditioned

auditory stimulus lasted 19 seconds. During the last ten seconds both stimuli were present. The interval between trials was varied between 10 and 30 seconds. Since the pupil quickly became so variable that changes due to the light were almost indistinguishable from the continual oscillations in size, it was found necessary to give the subject a rest between every three or four trials. The rest periods varied in length from 2 to 5 min. depending upon the desires of the subject.

The general procedure was as follows. At the first trial of the first session, the auditory stimulus was set at threshold or below threshold intensity (with the exception of subject We). After four to ten simultaneous presentations of light and sound, the sound was increased in intensity by two decibels, and further trials were given. This process was continued until the tone was well above threshold. It was felt that if conditioning were to appear in one or two trials at or near the threshold as claimed by Baker, the above procedure would afford the best conditions for its appearance at the most effective intensity.

It might be well at this point to summarize and compare our apparatus and method with that of Baker. In the first place, the levels of illumination at the eye of the subject were the same as those used by Baker. The stimulation times were the same, and so was the fixation distance. The auditory stimulus (CS) was delivered to the subject through headphones as was Baker's but the stimulus itself was an improvement over his, in that it was produced by a standard instrument and its intensity is defined in decibels so that it can be easily reproduced. The methods of measuring and recording changes in pupillary diameter are essentially the same although we state clearly the magnifying power of our recording system, which Baker does not do. The general disposition of our apparatus was an improvement over that of Baker because all apparatus which made any operating noise was located outside the soundproof room. All of Baker's apparatus was in one room and there was evidently considerable noise from his clock, time markers, audiometer motor, etc. In fact he states (1, p. 15) that even the supraliminal intensities used were not great as ". . . attested by the fact that some S's had initial difficulty in discriminating out this sound [of the CS] as compared with the low hum of the audiometer motor and the click of the Harvard clock." Finally, it should be stated that our method of limen determination did differ from that of Baker. However, since the intensity of the CS which we used ranged from a sound that was definitely below threshold to one that was definitely above, and since eight or more trials were given at each intensity level, this difference in method cannot account for the difference between our results and Baker's.

Thus Baker's method was duplicated as far as possible with every subject except 'We.' (No liminal nor subliminal tones were used with this subject.) However, not only were liminal or near-liminal auditory intensities used, but also various supraliminal intensities, in an attempt to obtain some conditioned responses. Consequently the number of trials at each level of intensity of the CS, and the number of test trials (sound without light) were varied from subject to subject. The number of test trials per subject was not great because we were looking for an anticipatory rather than a simultaneous CR.

Rather than discuss in detail the program for each subject the pertinent facts have been summarized in Table 1. The successive columns of the table show: 1. The letters designating the subject; 2. The direction of conditioning (dilation or constriction), and the conditioned stimulus (buzzer or 1024 cycle tone); 3. The number of the experimental session; 4. The interval between any session and the preceding session in days; 5. The intensity of the conditioned stimulus in decibels above or below the threshold at the beginning of the session; 6. The same at the end of the session; 7. The

number of combinations of CS and UCS at each session; 8. The number of rest periods in each session; 9. The number of test trials (CS without UCS); and 10. The number of responses which might be considered to be conditioned responses (quasi CR's). This last column is discussed in detail later. Table 2 is a recapitulation of Table 1 with the addition of a last column in which the number of quasi CR's is represented as a percent of the total number of trials for each subject.

TABLE 1

SUMMARY OF THE TREATMENT OF EACH SUBJECT IN THE EXPERIMENTAL GROUP

Subject	Direction of Conditioning and CS	Session	Interval between Sessions in Days	Intensity of CS in db Above or Below (−) Threshold		Number of Trials at Each Session	Number of Rest Periods	Number of Test Trials	Number of Quasi CR's
				At Start of Session	At End of Session				
Al	Dilation	1		−5	1	35	8	0	5
	1024 cycles	2	1	3	11	20	4	0	0
Wi	Dilation	1		−5	−1	12	2	0	3
	1024 cycles	2	1	−5	11	36	8	0	5
La	Dilation	1		−5	−3	16	3	0	2
	1024 cycles	2	1	−3	−1	15	2	0	0
		3	1	−1	3	24	4	0	8
		4	1	3	5	24	4	0	3
		5	1	3	7	30	5	0	2
		6	2	5	5	24	4	0	10
		7	1	5	5	14	2	0	2
		8	18	5	5	12	2	0	2
Fr	Constriction	1		0	2	18	5	0	0
	1024 cycles	2	1	2	4	18	5	0	2
		3	4	6	6	12	3	0	2
		4	1	8	8	13	3	0	0
		5	1	8	8	18	4	0	1
		6	2	8	8	14	3	0	2
		7	2	8	89	20	3	0	1
		8	1	89	89	12	2	0	3
		9	14	89	89	9	2	0	1
		10	3	89	89	15	4	0	2
		11	1	89	89	14	4	0	1
		12	17	89	89	20	4	0	2
		13	11	89	89	26	5	0	3
We	Constriction	1		10	10	25	4	0	2
	1024 cycles	2	1	10	20	64	12	1	9
Du	Constriction	1		−5	3	29	5	1	4
	1024 cycles	2	2	5	7	51	12	4	6
		3	1	7	7	36	7	0	5
		4	1	7	7	41	12	0	4
		5	5	7	7	30	8	1	8
		6	1	7	7	15	3	0	0
		7	12	7	7	33	7	0	11
		8	14	17	17	18	6	0	4
		9	2	17	17	11	2	0	1

TABLE 1—*Continued*

Subject	Direction of Conditioning and CS	Session	Interval between Sessions in Days	Intensity of CS in db Above or Below (−) Threshold		Number of Trials at Each Session	Number of Rest Periods	Number of Test Trials	Number of Quasi CR's
				At Start of Session	At End of Session				
Ru	Constriction	1		−2	−1	23	1	0	2
	1024 cycles	2	35	−2	4	24	4	0	4
		3	20	4	6	25	4	0	4
		4	23	6	6	20	4	0	2
Co	Dilation	1		−5	1	24	7	0	2
	buzzer	2	1	3	9	24	7	1	1
		3	1	11	17	24	7	0	2
		4	1	19	25	24	7	0	0
		5	1	27	33	24	7	3	5
		6	1	33	33	24	7	1	0
		7	2	33	33	24	7	2	0
		8	1	33	33	32	10	0	1
Br	Dilation	1		−5	−1	19	5	0	0
	buzzer	2	1	1	5	21	6	0	2
		3	2	7	13	24	7	0	0
		4	1	15	21	24	7	2	0
		5	1	23	29	24	7	0	0
		6	2	31	35	24	7	1	0
		7	1	35	35	24	7	1	1
		8	2	35	35	32	10	1	1
Wa	Constriction	1		−5	−1	18	5	0	5
	buzzer	2	12	−1	7	22	4	0	1
		3	0	9	15	20	3	0	3
Ot	Constriction	1		−5	−1	24	5	0	2
	buzzer	2	18	1	5	24	5	0	2
		3	0	7	11	24	5	0	5

It is not possible to determine from Table 1 exactly how many trials each subject had at each intensity of tone, but it is possible to do so with a small error by making use of the fact that the tone was usually increased in intensity in steps of two decibels. For example, Table 1 shows that subject Ot had 24 trials at the first session. At the beginning of the session the tone was 5 db below threshold, while at the end it was 1 db below the limen. The intensities used must therefore have been 5, 3, and 1 db below threshold with 8 trials at each intensity.

There were 11 subjects in all. Three were members of the staff, five were graduate students, and three were naive subjects who showed by their reactions that they remained ignorant of the purpose of the experiment until it was over and had been explained to them.

Results

Behavior of the Pupil.—Before discussing the evidence for conditioning, some observations on the general behavior of

the pupil seem worth mentioning. Measurements were made of the maximum response of the pupil to the unconditioned stimulus during the first five and the last five trials of each session for each subject. The average maximum constriction for six subjects was 3.0 mm. For the five subjects trained in the direction of dilation, the average maximum dilation was 2.3 mm. The largest single response to the UCS was 5.1 mm by subject Co, and the smallest was .8 mm by subject Br. These values are much smaller than those reported by Baker (1, p. 12). He reports an average maximum response of over 6 mm.[1] Our curves of the pupillary response (see Fig. 3) showed the typical differences between dilation and constriction. Constriction is very fast and takes place in less than a second. It is difficult to follow with the pupillometer. Dilation, on the other hand, is slow enough so that it is easy to follow, and it has seldom reached its maximum in the ten seconds during which the unconditioned stimulus is acting. These results are in agreement with the statements in the physiological literature that constriction is roughly 60 times as fast as dilation.

It must be remembered that the pupil is rarely quiet even when the light is constant, although the pupils of some subjects are very constant while others are not. At the beginning of each session the record of most subjects is characterized by many spontaneous oscillations in pupil size. With continued fixation and stimulation, the oscillations become greater and greater until they almost obscure the changes due to the light, and would most certainly obscure the conditioned response if it were present. Therefore it is necessary to give frequent rest periods as mentioned above.

Evidence for Conditioning.—It was immediately evident from a casual inspection of the records, that there was no unequivocal evidence for conditioning. By unequivocal evidence is meant an anticipatory response that persists for several trials, and is distinguishable from the above-mentioned normal oscillations in pupil size.

[1] Comparison of our values with those reported in the physiological literature shows that Baker has made an error in reporting his amplitudes of response. The values he gives should be divided by two.

Although there were no responses like the UCR and at the same time unlike the normal oscillations in pupil size, nevertheless there were many responses occurring during the nine-

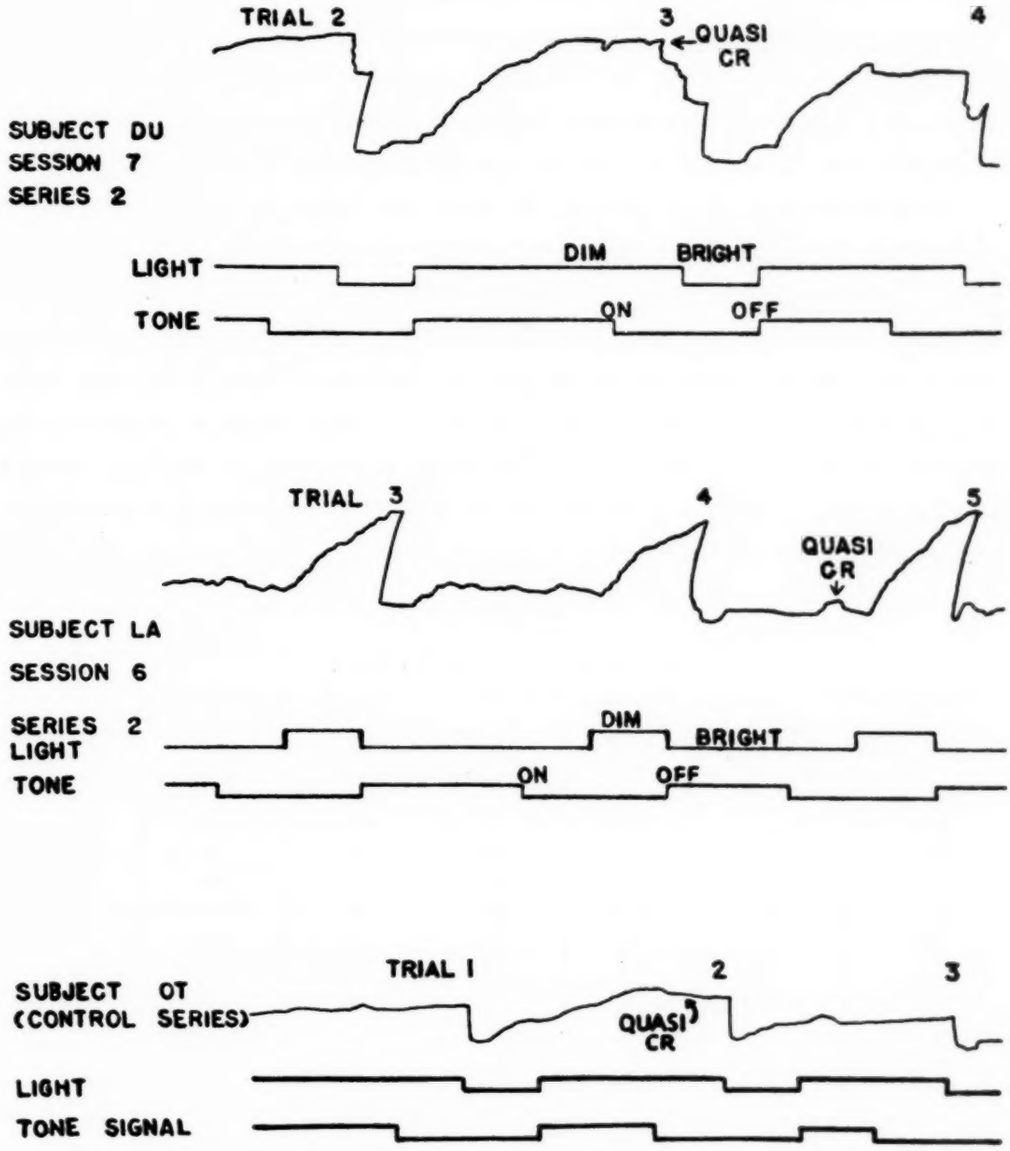


FIG. 3. Three sample records. A downward movement of the recording pen indicates constriction while an upward movement indicates dilation of the pupil. The time-line which would normally appear at the top of each record has been deleted. See text for discussion of these records.

second period when only the CS was present. Should these be called conditioned responses? What criteria can be used to distinguish the CR from other responses? Examination of the experimental literature shows that no single criterion can

be used. The latency of the CR is not the same as the latency of the UCR. The amplitudes of the CR and UCR are almost never the same. Finally, and most important, the form of the CR is not necessarily the same as the form of the unconditioned response. In fact, according to Hilgard and Marquis (**5**, p. 27) "Any new or altered response to the conditioned stimulus whose occurrence depends upon the repeated double stimulation is called a conditioned response."

Considering these facts, it was felt best to be on the safe side and consider any response occurring in the interval when the CS was operating alone to be a quasi CR, provided it was not obviously a normal fluctuation in pupil size. Consequently, the records were inspected independently by the two senior authors. Whenever the two agreed that a certain response might be called a conditioned response, note was made of this fact. The last column in Table 1 and the last two columns in Table 2, show the frequency and distribution of these

TABLE 2

RECAPITULATION OF TABLE 1

Subject	Total Number of Sessions	Total Trials	Total Rest Periods	Total Test Trials	Total Quasi CR's	Quasi CR's as % of Total Trials
Al	2	55	12	0	5	9
Wi	2	48	10	0	8	17
La	8	159	26	0	29	18
Fr	13	209	47	0	20	10
We	2	89	16	1	11	12
Du	9	264	62	6	43	16
Ru	4	92	13	0	12	13
Co	8	200	59	7	11	6
Br	8	192	56	5	4	2
Wa	3	60	12	0	9	15
Ot	3	72	15	0	9	13

quasi CR's. By designating them as quasi conditioned responses, the authors mean to indicate that they do not believe they are conditioned responses but cannot prove that other experimenters would have the same opinion. It should be emphasized that in no case was the amplitude of the response as great as that of the CR's obtained by Baker (**1**, p. 10ff.). He reports the CR to be two-thirds the size of the UCR after the third training combination.

An inspection of the last column of Table 2 shows that the percentages of quasi CR's range from 2 to 18 with a mean of 12. It may be noted that in some sessions with certain subjects the number of quasi CR's is large. For example, Table 1 shows that in session 6, subject La gave 10 responses out of a total of 24 trials. Their occurrence in succession would certainly be evidence for conditioning even though the response disappeared in subsequent trials. However the greatest clustering of these responses was three times in succession with four subjects. They appeared twice in groups of three each with subject La, and once in clusters of the same size with each of two other subjects. In all other instances they were scattered throughout the session. Nor did they always appear at the beginning of a session nor at the end nor at any other constant position in the series.

Control Experiment.—In spite of the lack of clustering of the quasi conditioned responses, their number is considerable. Therefore it is necessary to show that the observed frequencies are no greater than would be expected by chance and are not due to the repeated double stimulation with tone and light. In other words, some way must be found to answer the question, with what frequency must a response to a CS occur in order to be considered a CR? If it can be shown that the frequency of these responses in a control group of subjects treated in exactly the same manner as the experimental group except that the auditory stimulus is omitted is as great as in the experimental group, then the conclusion must be drawn that the observed frequency in the experimental group is not significant.

Accordingly a control experiment was run using eight subjects. Three of the eight were subjects who had been used in the original experiment, two were graduate students, and three were undergraduates who were ignorant of the purpose of the experiment. Three of the subjects were trained for dilation and five for constriction.

In the control experiment, conditions were precisely the same as in the original experiment except that no tone or other auditory stimulus was present. The subjects even wore the

headphones and were led to believe that there might be a subliminal tone present. All eight subjects were given 50 trials distributed between two sessions, usually with 30 trials in the first session and 20 in the second. The records were again inspected independently by the two senior authors. The method of inspection was to observe carefully the behavior of the pupil during the nine second interval previous to the change in light intensity (when the tone would have been present) and to note any responses which would have been called conditioned responses in the original experiment. If both agreed on a response, it was called a quasi CR.

An inspection of the records shown in Fig. 3 will illustrate most of the problems involved in determining whether a CR is present and will give a clue as to the character of the responses which were called quasi CR's. Portions of the records of three subjects are shown. The record of subject Du shows a quasi CR in trial 3 as noted in the Figure. There are also responses in trials 2 and 4 which some experimenters might have considered conditioned responses. The middle record in Fig. 3 is from subject La. A quasi CR is shown in trial 5. Notice that although this response was judged to be a possible or quasi CR, it is scarcely larger in magnitude and not essentially different in form from oscillations in pupil size occurring during the time when the tone alone was present in the two preceding trials, 3 and 4. The bottom record in Fig. 3 differs from the two above in that no tone was present. It is a sample record from subject Ot during the control experiment when no CS was present although the tone-pen was operating as usual. Here again a quasi CR is shown in trial 2. The absolute amplitude of this quasi CR is small, but relative to the size of the UCR, it is as large as those in the above two records. In other words, a constriction of the pupil occurred which would have been called a quasi CR had the tone been present.

In terms of percent of total trials, the frequencies of quasi CR's in the control group were found to be just about the same as in the experimental group. The percentages are: 4, 8, 10, 12, 14, 14, 16, and 20. The three subjects who were used in

both experiments were Du, Ot and Wa. In the control experiment they gave the following percentages of quasi CR's: 14, 16 and 8. The corresponding percentages in the original training situation were 16, 13, and 15 as can be seen in Table 2. Considering the small number of subjects, the agreement between the experimental and control situation is remarkably good. The results obtained in the control experiment show clearly that the responses obtained from the first group cannot be considered to be due to the combined presentation of tone and light. They are not conditioned responses to the tone. However, there is another possibility. What we have called quasi conditioned responses may be due to a kind of temporal conditioning. It may be that the subjects in both the control and experimental groups learned to expect the light within a certain variable interval. In any given trial, expectation of the UCS might cause a change in pupil size. However, since the quasi responses looked exactly like the spontaneous pupillary changes, and since the subjects were not apprehensive of the change in light intensity, the probability of temporal conditioning seems very slight indeed. Furthermore, the question of temporal conditioning as the explanation of the results in both experimental and control groups is not within the scope of this paper.

Discussion

Two conclusions are drawn by Baker which are important for the theory of conditioning. The first (1, p. 31) is that conditioning to subliminal auditory stimuli can be established in two trials. The second is that conditioning to subliminal auditory stimuli is more easily established than to supraliminal stimuli. These conclusions are not supported by the present experiment. Conditioned pupillary responses to light were not obtained under any conditions much less in one or two trials even though many levels of auditory intensity both above and below the limen were employed.

It may be objected that since the present experiment is entirely negative, the possibility that such conditioning as Baker's may take place under certain conditions has not been

eliminated. This is true. What the present experiment does show is that a thoroughgoing duplication of the conditions of Baker's experiment was entirely unsuccessful in producing his results. The conclusion must be drawn that the conditions stated by Baker may be necessary but they are not sufficient. Some other factors, as yet undetermined, must be important in controlling the formation of the conditioned pupillary response.

Finally, there are some suggestions which may be of interest to others. The first is that subjects should be selected with a view to making it as easy as possible to determine whether conditioning has taken place. The subjects should have light-colored irises so that the boundary of the pupil is clearly defined. They should also have pupils which remain relatively constant in size with constant experimental conditions, and yet show a large variation in pupil size in response to a change in the intensity of the unconditioned stimulus. Some subjects have pupils which show remarkably little variation in pupil size even with great variations in the intensity of the light.

A second consideration which seems important is that the stimulus times were too long. Since the conditioned stimulus lasted 19 seconds, the subjects became fatigued, and the whole experiment was unduly attenuated. Thirdly, the fixation point was too near the subject. Seven and one-half inches is not a comfortable fixation distance. Lastly instead of maintaining the time between trials within certain limits (between 10 and 30 seconds in this experiment), it would seem favorable to an easy analysis of the records to stimulate at any time when the pupil is quiet, and to wait until the pupil is quiet before stimulating.

Summary

An attempt was made to condition the change in size of the pupil in response to changes in the intensity of light to a tone of 1024 cycles or the sound of a buzzer. Many subliminal and supraliminal levels of auditory intensity were employed. No conditioning was obtained although the conditions under which a previously successful attempt was made by another

experimenter were duplicated as far as possible. It is pointed out that these conditions were not ideal for the establishment of a conditioned response, and that the failure means that some unreported or unknown factors must have contributed to the success of the previous attempt.

(Manuscript received June 17, 1940)

REFERENCES

1. BAKER, L. E., The pupillary response conditioned to subliminal auditory stimuli, *Psychol. Monog.*, 1938, **50**, No. 3, 32 pp.
2. CASON, H., The conditioned pupillary reaction, *J. exper. Psychol.*, 1922, **5**, 108–146.
3. FERREE, C. E., AND RAND, G., An instrument for measuring the breadth of the pupil, *Amer. J. Psychol.*, 1927, **38**, 292–293.
4. HARLOW, H. F., AND STAGNER, R., Effect of complete striate muscle paralysis upon the learning process, *J. exper. Psychol.*, 1933, **16**, 283–294.
5. HILGARD, E. R., AND MARQUIS, D. G., *Conditioning and learning*, N. Y.: D. Appleton Century, 1940.
6. HUDGINS, C. V., Conditioning and the voluntary control of the pupillary light reflex, *J. gen. Psychol.*, 1933, **8**, 3–51.
7. METZNER, C. A., AND BAKER, L. E., The pupillary response conditioned to subliminal auditory stimuli: a control experiment, *Psychol. Bull.*, 1939, **36**, 625.
8. STECKLE, L. C., Two additional attempts to condition the pupillary reflex, *J. gen. Psychol.*, 1936, **15**, 369–377.
9. STECKLE, L. C., AND RENSHAW, S., An investigation of the conditioned iridic reflex, *J. gen. Psychol.*, 1934, **11**, 3–23.
10. WATSON, J. B., The place of the conditioned reflex in psychology, *Psychol. Rev.*, 1916, **23**, 89–116.

RETROACTIVE AND PROACTIVE INHIBITION IN IMMEDIATE MEMORY *

BY W. B. PILLSBURY AND ANNE SYLVESTER

The investigation was undertaken to determine whether retroactive inhibition could be detected after one repetition of items and immediately after the presentation. The great mass of earlier work deals only with material learned to one perfect repetition and is tested after several hours or a day. We gave but one repetition and tested retention after ten seconds, which was just sufficient for the work that was to produce retroaction. Later we compared the effect of proaction with retroaction by asking the observers to observe series of the same kind of material and be tested for it, just before seeing the experimental series.

Two methods of presentation were used. In the first series the items were presented by a tachistoscope, eight at an exposure. In the other series six items were used and they were presented serially by a modified Spindler and Hoyer memory apparatus. The items presented were geometrical figures, three letter words, and nonsense syllables. The geometrical figures were squares, triangles, circles and crosses and the test was to give the order in which they were arranged. The observer sat at a table about two feet in front of the tachistoscope or the memory apparatus. Before him was a sheet of paper, ruled with as many lines as there were experiments. This sheet was covered by another which could be moved down to expose one line of the problem sheet at a time. If the experiment was a control, the line was blank. When work was to be done the task was given on the line. Arithmetic problems were given and questions were asked with a space left for the answer.

As the exposure was made on the tachistoscope, the sheet

* This investigation was aided by grants from the Horace H. Rackham and Faculty Research Funds of the University of Michigan and by aids from the National Youth Administration.

was drawn down to expose one line. If it were a retroaction experiment, the observer began at once to write the answer to the problem and continued for ten seconds when the signal to stop was given. He then wrote what he had seen. In the control experiments, as was said above, he waited ten seconds and then wrote what he had seen.

TABLE I

	Words No Retro.	Words Arith. Retro.	Words Written Retro.	Geom. Fig. No Retro.	Geom. Fig. Arith. Retro.	Geom. Fig. Written Retro.
K.						
Mean.....	4.5	2.19	1.67	4.34	2.31	2.20
S.D.....	1.12	1.38	1.42	1.24	1.39	1.52
N.....	94	161	65	90	162	62
T.						
Mean.....	4.34	3.94	3.30	4.57	4.08	3.71
S.D.....	.80	.98	1.29	.60	1.04	1.27
N.....	93	165	62	97	160	64
S.						
Mean.....	3.58	2.46	1.97	4.37	2.42	2.42
S.D.....	.96	1.17	1.11	1.01	1.48	1.38
N.....	82	165	63	98	152	62
F.						
Mean.....	3.84	3.54	2.81	4.04	3.61	3.32
S.D.....	.74	.78	1.14	.92	1.02	1.29
N.....	89	164	62	86	159	65
Har.						
Mean.....	3.87	3.21	2.03	3.81	3.41	2.22
S.D.....	.64	1.01	1.09	1.00	1.18	1.53
N.....	90	166	62	90	158	64
Hor.						
Mean.....	4.42	3.18	2.22	4.72	2.68	2.71
S.D.....	1.03	1.53	1.48	1.22	1.06	1.54
N.....	92	157	62	89	154	58

The results are shown in Table I. The first column gives the number of words that could be reported when presentation was followed by a rest. The total number shown in each case was eight. The second column gives the number of items that were recalled when arithmetic problems were solved immediately after the exposure. In the third column are given the numbers that could be recalled after answering questions for ten seconds. In the last three columns are the same results

when the material is geometrical figures, the fourth gives the values when the presentation is followed by a rest, the fifth when followed by arithmetical problems and the sixth when followed by answering questions. It will be seen that the decrease in the number of items reported when exposure is followed by work as compared with the times when it is followed by a rest is great in each case. For some observers the number reported is decreased by nearly 50 percent, in all but one by 10 percent or more. Observer F., for whom the decrease is least, paid little attention to the tasks and accomplished less than the others. K. frequently finished all tasks in the ten seconds and was most affected. The standard deviations are large, sometimes larger than the differences. Nevertheless the consistency of the results for all of the individual observers gives them significance. In no one of the six observers is the number of items reported larger after work than after rest. It may be noted that the standard deviations are also larger after work than after rest.

In a second series of experiments the same kind of material was exposed serially with the memory apparatus. The separate items were printed on slips of paper and exposed for a quarter of a second each. The six items of each series required three seconds for the showing. The method used in the first series for introducing work and recording the results was used again here. A slight rest was introduced between the series. After the fifth series in this experiment we measured the times required for the response. A stop watch measured the time taken from giving the experiment to the cessation of writing the report. The results, including the times, are given in Table II.

It will be seen that the results in this series are very similar to those of the first experiment. With a single exception each test on all of the nine subjects markedly reduces the number of items reported. The one exception, J. for geometrical figures, shows a marked increase in the time of response. On the whole the increase in time is quite as regular as was the decrease in the number of items reported for the series in which presentation was followed by work. The number of

TABLE II

	Geom. Fig. No Retro.	Geom. Fig. Arith. Retro.	Geom. Fig. Written Retro.	Words No Retro.	Words Arith. Retro.	Words Written Retro.	Non. Syl. No Retro.	Non. Syl. Arith. Retro.	Non. Syl. Written Retro.
K.									
Mean	4.81	2.82	2.78	5.03	2.79	1.58	2.62	.79	.67
S.D.	1.50	1.61	1.48	1.31	1.38	1.20	1.23	.77	.71
M. Time	4.46	5.37	5.62	6.68	5.42	7.54	6.08	5.11	5.50
S.D.	1.52	1.38	1.74	1.31	1.91	2.84	1.22	1.55	1.09
N	32	48	32	32	48	32	32	48	32
H.									
Mean	4.54	4.04	2.87	4.00	3.41	2.84	2.31	1.49	1.31
S.D.	1.30	1.26	1.30	1.42	1.07	1.38	1.28	.98	1.13
M. Time	4.21	5.36	5.21	5.75	5.91	5.50	5.00	4.80	5.54
S.D.	.91	1.24	1.32	1.27	1.48	1.35	1.13	.71	1.38
N	32	48	32	32	48	32	32	48	32
R.									
Mean	4.25	3.10	2.68	4.62	3.36	2.32	1.68	.75	.49
S.D.	1.50	1.88	1.77	1.13	1.21	1.44	.88	.63	.79
M. Time	6.71	12.22	11.96	10.66	15.60	18.50	13.08	14.94	17.29
S.D.	2.73	3.35	4.20	3.46	3.65	4.36	3.68	4.05	5.02
N	32	48	32	32	48	32	32	48	32
M.									
Mean	5.40	4.19	4.97	5.66	3.87	3.80	2.63	1.07	1.14
S.D.	1.22	1.72	1.02	.47	1.52	1.58	.91	1.03	.83
M. Time	5.83	12.03	13.75	9.75	21.94	17.84	16.58	20.18	20.58
S.D.	1.74	4.20	4.38	4.01	4.53	5.09	4.30	4.20	4.09
N	32	48	32	32	48	32	32	48	32
V.									
Mean	4.22	3.06	2.78	5.47	3.91	3.06	1.80	.83	.58
S.D.	1.60	1.59	1.76	.77	1.23	1.79	.94	.66	.65
M. Time	7.37	8.64	10.54	9.08	10.71	12.50	7.41	9.25	10.45
S.D.	2.21	2.71	2.64	1.86	2.40	4.09	2.24	3.12	3.40
N	32	48	32	32	48	32	32	48	32
J.									
Mean	3.21	3.55	3.19	4.00	3.64	2.34	2.00	1.33	1.00
S.D.	1.50	1.59	2.10	1.07	1.33	1.16	.71	.76	.83
M. Time	6.80	8.11	8.00	9.75	9.33	10.25	8.08	6.66	9.16
S.D.	2.24	2.10	3.09	2.70	1.75	1.80	3.20	2.70	3.22
N	32	48	32	32	48	32	32	48	32
B.									
Mean	5.03	4.16	3.96	5.71	3.54	3.06	3.08	1.64	1.66
S.D.	1.20	1.49	1.71	.56	1.33	1.33	1.08	1.01	.73
M. Time	10.00	15.20	15.44	11.24	12.55	26.25	18.29	24.11	29.34
S.D.	3.5	4.15	5.00	3.35	4.40	6.15	5.20	5.00	3.10
N	32	48	32	32	48	32	32	48	32
T.									
Mean	2.72	2.43	2.19	3.71	1.85	1.17	1.03	.64	.47
S.D.	1.68	1.36	1.50	1.21	1.37	.95	.78	.68	.62
M. Time	12.80	17.41	19.29	18.50	24.11	25.96	19.17	21.58	21.33
S.D.	5.09	3.75	4.60	4.40	4.20	4.30	4.70	3.01	3.60
N	32	48	32	32	48	32	32	48	32

nonsense syllables reported was considerably smaller than the number of short words or geometrical figures. Retroaction showed a greater effect also when nonsense syllables were to be recalled. Again solving general problems produced a greater decrease in the number of items recalled than did the arithmetical problems.

While both types of work in these two experiments were different from the items to be recalled, the decrease in number retained for ten seconds amounted to from about 10 percent to 70 percent of the number retained when the ten seconds was spent in rest. Neither form of work performed in this series had any resemblance to remembering and the items involved in the work were not similar to the items remembered.

This left open the question as to whether retroactive inhibition was dependent at all upon associative inhibition or the confusion of items remembered in the test series with items in the retroaction series. To test this hypothesis a third experiment was carried out in which the work done to produce the retroaction was in each case learning nonsense syllables, while the original material included geometrical figures, short words, and nonsense syllables as before. After the presentation of the list of items that were to be recalled, the experimenter read the observer a list of six nonsense syllables and he was asked to write as many of them as he could remember. At the end of ten seconds he was asked to write as many of the original series as he could remember. Seven series of each type of material was shown at each test. Four of these series were followed by learning nonsense syllables and three of each were followed by a rest.

The results are shown in Table III. The retroactive effects of learning nonsense syllables are much greater than of solving arithmetical problems or of answering general questions. It is shown both by the smaller number of items recalled and by the longer times taken for the recall. Reporting on nonsense syllables reduced least the number of geometrical figures recalled, most a series of other nonsense syllables, while the effect upon the retention of short words was intermediate.

TABLE III

	Geom. Fig. No Retro.	Geom. Fig. Non. Syl. Retro.	Words No Retro.	Words Non. Syl. Retro.	Non. Syl. Non. Syl. Retro.	Non. Syl. Non. Syl. Retro.
K.						
Mean.....	4.49	2.40	5.36	1.75	2.60	.42
S.D.....	1.31	1.51	.89	1.20	.88	.76
M. Time....	5.20	6.00	7.60	6.93	5.56	5.47
S.D.....	1.67	1.70	1.86	2.69	1.58	1.85
N.....	30	40	30	40	30	40
H.						
Mean.....	4.40	3.00	4.37	2.47	2.59	1.22
S.D.....	1.22	1.36	1.02	1.11	.79	.79
M. Time....	5.26	4.72	6.33	5.50	5.56	5.75
S.D.....	1.22	1.08	1.30	1.04	1.27	1.47
N.....	27	36	27	36	27	36
T.						
Mean.....	3.75	1.86	3.56	1.16	1.31	.42
S.D.....	1.83	1.27	.95	1.04	.97	.55
M. Time....	9.44	15.31	15.63	19.63	11.80	16.18
S.D.....	3.50	3.79	4.50	4.90	3.69	4.30
N.....	35	35	35	35	35	35
R.						
Mean.....	4.66	1.07	5.17	.92	2.67	.40
S.D.....	1.65	1.22	.80	.75	.92	.48
M. Time....	8.33	16.94	15.00	20.45	17.65	20.23
S.D.....	3.75	4.47	6.49	4.80	3.80	4.39
N.....	35	35	35	35	35	35
B.						
Mean.....	5.73	2.87	5.26	1.73	2.86	1.23
S.D.....	.62	1.75	.98	1.39	1.08	.91
M. Time....	8.63	22.75	15.86	24.88	15.03	23.60
S.D.....	3.61	7.58	5.95	5.57	4.85	6.75
N.....	30	40	30	40	30	40
J.						
Mean.....	4.20	3.29	4.00	1.78	2.30	.58
S.D.....	1.19	1.89	1.16	1.17	.90	.73
M. Time....	7.06	9.35	11.00	11.15	8.40	10.21
S.D.....	1.96	1.76	2.43	3.38	2.64	2.83
N.....	30	40	30	40	30	40
M.						
Mean.....	5.67	3.64	5.70	2.98	3.07	1.08
S.D.....	.20	1.43	.46	1.53	.91	1.03
M. Time....	5.53	16.25	10.80	21.01	21.90	23.32
S.D.....	1.52	4.90	3.12	5.30	4.45	5.39
N.....	30	40	30	40	30	40

Since this experiment showed that learning nonsense syllables was rather more effective than doing work other than learning, it was decided to compare the retroactive effect of learning each type of material upon the retention. In Ex-

periment IV each set of six geometrical figures was followed by six other geometrical figures, each set of six words by six other words and each set of six nonsense syllables by six other nonsense syllables. The results are given in Table IV. The

TABLE IV

Observ.	Geom. Fig. No Retro.	Geom. Fig. G. F. Retro.	Words No Retro.	Words W. Retro.	Non. Syl. No Retro.	Non. Syl. N. S. Retro.
W.						
Mean	4.38	3.00	4.43	2.50	3.03	1.15
St. Dev.	1.21	1.95	1.05	1.50	.94	1.00
Mn. Time	6.18	9.68	9.07	13.70	9.20	11.90
St. Dv. T.	1.43	1.80	1.78	2.33	1.15	2.05
No.	30	40	30	40	30	40
B.						
Mean	3.00	1.43	4.90	.96	2.06	.25
St. Dev.	1.00	1.01	1.37	1.02	1.08	.43
Time	8.50	11.00	13.20	10.00	13.00	13.60
St. Dv. T.	1.45	1.93	1.94	2.70	2.51	2.57
No.	30	40	30	40	30	40
J.						
Mean	3.93	2.07	3.67	1.21	2.26	.48
St. Dev.	1.40	1.33	.50	.94	.83	.46
Time	6.43	6.60	6.80	7.20	8.47	8.30
St. Dv. T.	1.41	1.47	.97	1.37	1.59	1.73
No.	30	40	30	40	30	40
C.						
Mean	4.25	2.53	4.90	1.11	3.00	.28
St. Dev.	1.25	1.28	1.04	.83	1.18	.50
Time	9.10	11.70	13.00	20.07	13.10	18.00
St. Dv. T.	2.36	3.05	2.91	2.70	2.35	3.88
No.	30	40	30	40	30	40
P.						
Mean	5.43	2.37	4.53	1.84	2.56	.40
St. Dev.	.91	2.07	1.41	.91	.74	.48
Time	6.28	7.70	9.28	7.70	7.97	5.43
St. Dv. T.	1.37	1.59	1.94	2.04	1.97	1.31
No.	30	40	30	40	30	40
E.						
Mean	4.30	1.43	4.60	1.90	2.13	.56
St. Dev.	1.58	1.17	1.04	1.24	.81	.70
Time	7.70	9.40	13.20	13.20	11.80	11.25
St. Dv. T.	1.51	1.80	1.97	1.83	2.16	2.58
No.	30	40	30	40	30	40
O.						
Mean	3.51	2.27	4.03	2.41	2.43	1.00
St. Dev.	1.28	1.09	.99	.98	.72	.68
Time	9.28	12.00	8.35	10.64	8.22	12.90
St. Dv. T.	1.57	2.91	1.45	1.78	1.99	2.57
No.	30	40	30	40	30	40

number of items reported when the first repetition is followed by rest is given for geometrical figures in column 1, for short words in column 3, and for nonsense syllables in column 5. Columns 2, 4, 6, give the number reported when the presentation is followed by presenting and testing recall for six other items of the same sort before the test of the first series is made.

The observers were not the same as in the third experiment but were of nearly the same learning ability. A comparison of the results in Tables III and IV offers some indication of the relative effect of the same and different types of material in disturbing the number of items recalled. In Table III, retroaction was due to seeing and reporting on nonsense syllables. In Table IV retroaction was due to seeing and reporting on material of the same kind as in the first series. We averaged the items recalled by the eight individuals and determined the percentage reduction when report was preceded by learning nonsense syllables over the decrease during a rest period. The decrease was greater by 55 percent for geometrical figures, by 62 percent for short words and 77 percent for nonsense syllables. When each type of material was followed by a report on another series of the same type the increased reductions were 45 percent, 62 percent, and 69 percent. Since nonsense syllables were followed by nonsense syllables in both experiments there would be a significant comparison for two cases only, geometrical figures and short words. Reporting on nonsense syllables probably required more work than learning short words or geometrical figures. At least fewer are reported by each observer in both series of tests. Recalling them produces a greater retroactive effect than recalling geometrical figures after geometrical figures, 55 percent against 45 percent. For short words recalling other words had exactly the same effect as recalling nonsense syllables, 62 percent. On the whole, these results indicate that inhibition from similarity of material learned is less important than from the severity of mental preoccupation. In one case the report on nonsense syllables caused a greater reduction in the succeeding report on geometrical figures than did a report on other geometrical figures. For words, the

reduction was the same for both. The greater effort in attempting to recall nonsense syllables more than compensated for any increased effect of reporting on similar as compared with dissimilar material.

These results may be compared with those given in the first two tables where the work that followed presentation was answering questions or solving arithmetical problems. These tasks were probably less like the material to be reported on than recalling any material. Working arithmetical problems after seeing geometrical figures reduced the number recalled 25 percent more than with a rest period and reduced the number of short words recalled by 28 percent. Answering general questions reduced the number of geometrical figures reported on by 43 percent and the number of short words by 36 percent. In the second series, where the items were shown serially, working arithmetical problems reduced the number of geometrical figures reported by 20 percent and the number of short words by 34 percent. Answering general questions reduced the number of geometrical figures recalled by 26 percent and the number of short words by 47 percent. Students reported that answering questions was harder than solving arithmetic problems before they or we knew the results. There is probably little difference in the degree of similarity with the items to be recalled. Whether either work is more or less difficult than recalling other items cannot be determined.

The results show no clear cut evidence as to whether similarity or intensity of work after the presentation of the material is more important in reducing the number of items recalled. Where the work was learning, reviving the more similar material reduced the number recalled less than recalling the more unlike. Where there was little similarity, the more difficult task reduced the number more than did the easier. If we assume that learning is more difficult than either of the other tasks the retroactive effect would be proportional to the difficulty. There would be no definite conclusion to be drawn as to similarity. Of the forms of learning the more similar produces the greater amount of retroaction, and learn-

ing shows a greater effect than either of the other tasks. No one doubts that there is a confusion between similar items, the basis of associative inhibition. These results do show, I believe, that work with little similarity to the items learned also produces retroaction.

The remarks of the observers on special phases of the work throw some interesting light on special aspects of the problem. They were all asked to refrain from thinking of the material that had been shown them during the rest period after the presentation. Many of them reported that they could not always be sure that they had complied with this request. Some said that the impressions would recur to them in spite of their efforts. However, those who were most sceptical of avoiding recall were those who showed the least difference between the work period and the rest period. It would imply that the conscious return did not appreciably influence the retention.

Some of the observers objected that the rest period of ten seconds might be a disadvantage. They said that if they were permitted to write at once they would recall much more. In the first series the comparison rest experiments were omitted and tests made at once after presentation. Then it seemed that the conditions were not comparable with the work series in which ten seconds elapsed between presentation and the test. The smaller number reported might have been due to the interval. Tests in the series proved the number reported after ten seconds was little different from the number reported immediately, and it was as frequently greater as smaller.

Some comments on the imagery used were suggestive. Some used visual imagery directly. Others reported that they translated the geometrical figures into words and remembered the words. On the whole these individuals were more accurate than the average in the material that they reported but they reported a smaller number. Several translated the first four only and these would be recalled correctly in spite of all the work that might be done. Individuals who retained the items visually would report a larger number, but would make more mistakes. The later learning affected the recall of these individuals more than of the others.

Fatigue needed to be avoided if the results were to be constant. A series occupied from 25 to 35 minutes and a ten minutes pause was given between series. When more than two series was attempted in an afternoon, the accuracy of the work fell off. In consequence no more than two series were given. Practice effect was slight. Improvement ceased at about the middle of the first test.

Discussion

These experiments differ from the traditional ones on retroaction in that they are applied to the results of primary memory and show that the reduction in retention takes effect in less than ten seconds after the trace is left. Either the Müller hypothesis or the more recent explanation in terms of associative inhibition may be applied to explain them. On the first hypothesis, the immediate trace would be wiped out by any mental work. On the second, forming a second association disturbs an original one. More generally, in learning, new items tend to become fused with earlier presented material and make recall more difficult. In our case there were no identical elements involved in successive series, hence true associative inhibition was not an explanation. Some confusion between different items seen might have been present. The fact cited above that learning nonsense syllables was more effective in one series and equally effective in another than learning other series of the same kind would reduce the probability of this explanation. Of course, retroactive inhibition has been fully demonstrated in other connections and may contribute to the difference here. Work of any kind will also greatly reduce the number of items that can be recalled.

In passing it is interesting to compare the results of work on retention of material presented earlier with the effect of performing a second task at the same time as the one to be tested. Distraction experiments from Dallenbach on show that interference with a task is likely to increase efficiency in performance. Work done after the impression has been made decreases its efficiency. The explanation may be that one prepares for the distraction by extra effort as Morgan and

Ford demonstrated, while there is no chance to prepare for the work that comes after.

Proactive Inhibition

Closely related to retroaction is the problem of proaction. Not only learning that occurs after any presentation diminishes the retention of what was learned before, but material learned before the presentation of the test material also has been shown to decrease the amount retained of the material learned afterwards. These results were obtained with material that was learned to one accurate repetition and was relearned after a day or more. Our experiment was done by presenting lists of six words, geometrical figures and nonsense syllables and asking the observer to give the list ten seconds after the presentation. In the test series a second list of six words, geometrical figures, and nonsense syllables was given just before the primary series. The observer was tested for this first series during a ten seconds interval, the primary series then was tested. In the comparison series there was a rest before the learning of the primary series. The results are given in Table V. It will be seen that proaction has a very decided effect. The reduction is in the neighborhood of 50 percent in most cases and in many instances more.

An opportunity is given to compare the effects of retroaction with proaction for four of the observers were used both in this experiment and in Table IV of the study on retroactive inhibition. The percentage diminution for retroaction and proaction is strikingly similar. For B. retroaction for geometrical figures is 52 percent, for words 80 percent, for nonsense syllables 88 percent, while the corresponding results for proaction are 53 percent, 90 percent, and 85 percent. For E. the corresponding values for retroaction are 67 percent, 60 percent, and 74 percent, while for proaction they are 47 percent, 70 percent and 64 percent. For J. retroaction reduces the number of items retained by 29 percent, by 67 percent and by 79 percent, while proaction reduced the number of items recalled by 30 percent, by 82 percent, and by 77 percent for the different sorts of items. For P. retroaction re-

duced the number of items recalled by 56 percent, by 57 percent, and by 85 percent, respectively, while proaction reduced them by 43 percent, by 54 percent and by 50 percent. It will

TABLE V

PROACTIVE INHIBITION

Observ.	Geom. Fig. Rest	Geom. Fig. Proact.	Words Rest	Words Proact.	Non. Syl. Rest	Non. Syl. Proact.
B.						
Mean	3.55	1.67	4.50	0.45	1.55	0.23
S.D.	1.48	1.11	1.38	0.55	0.70	0.38
Time	9.43	10.00	13.70	9.40	14.10	11.28
S.D.T.	2.10	1.55	1.95	1.27	1.62	2.00
N.	30	40	30	40	30	40
E.						
Mean	4.37	2.20	4.78	1.42	1.10	0.40
S.D.	1.31	1.20	0.97	0.93	0.84	0.49
Time	7.03	7.78	12.40	11.22	11.60	11.78
S.D.T.	1.42	1.85	2.01	1.53	1.90	1.68
N.	30	40	30	40	30	40
J.						
Mean	2.82	1.95	4.26	0.76	1.73	0.40
S.D.	1.04	1.17	0.90	0.76	1.05	0.50
Time	7.60	6.12	8.47	8.30	8.73	8.68
S.D.T.	1.57	1.11	1.12	0.95	1.01	1.20
N.	30	40	30	40	30	40
P.						
Mean	5.10	2.90	5.43	2.52	2.50	1.26
S.D.	0.88	1.57	0.84	1.06	0.85	1.01
Time	6.27	7.05	9.03	7.70	7.33	6.90
S.D.T.	0.93	1.61	1.48	1.61	1.56	1.24
N.	60	80	60	80	60	80
Sm.						
Mean	5.57	1.65	5.60	1.85	2.16	0.60
S.D.	0.44	1.42	0.41	1.04	1.12	0.67
Time	5.60	5.10	10.50	6.40	10.10	6.00
S.D.T.	0.74	1.81	1.56	1.67	1.85	1.54
N.	60	80	60	80	60	80
St.						
Mean	5.80	2.45	5.27	1.88	3.63	0.88
S.D.	0.31	0.84	0.69	0.94	0.66	0.70
Time	4.70	5.90	8.20	5.90	6.60	6.30
S.D.T.	0.91	1.14	1.26	1.56	1.62	1.63
N.	60	80	60	80	60	80

be seen that retroaction and proaction reduce the number of items recalled in much the same ratio. With a few exceptions each is more evident for nonsense syllables and words than for geometrical figures.

Retroaction and proaction would both be active in continuous learning of a long series. Together they can be regarded as accounting for the great increase in number of repetitions required for learning long as compared with short series. Each reduces the number of items reported after ten seconds by from 25 percent to 85 percent. If they are at all additive in their effect they would together explain the great increase in number of repetitions required for one series of twelve items as compared with learning two series of six items each.

Summary

This is a study of the effect of retroactive inhibition upon primary memory. It shows that the number of items reported when they are seen in a tachistoscope or exposed once serially in a memory apparatus is reduced by from 25 percent to 43 percent when mental work is done during the succeeding ten seconds as compared with the number reported after a ten second rest.

Reporting on a second list of material has rather more effect (45 percent to 85 percent) than does mental arithmetic. Learning nonsense material has about the same effect as does learning material of the same kind as in the original list.

Proaction was also found to affect the retention of material in primary memory in about the same amount as does retroaction.

(Manuscript received June 14, 1940)

References

1. Brenner, L. R., An experimental investigation of memory span, *J. exper. Psychol.*, 1940, **26**, 467–482.
2. Twining, P. E., The relative importance of intervening time and lapse of time in the production of forgetting, *J. exper. Psychol.*, 1940, **26**, 483–501.

AN EXPERIMENTAL STUDY OF VARIABILITY IN OCULAR LATENCY

BY ROY B. HACKMAN

Temple University[1]

Since the history of ocular reaction-time studies has been admirably summarized by Miles,[2] it will not be discussed here in any detail. Instead we will point out the salient findings. The early studies of Dodge showed that the reaction-time of the eye could be measured accurately by means of the photographic technique developed by him. The mean latency was found to be about 200 sigma with a considerable amount of variation under different conditions. This result was for untrained subjects who had no exact knowledge of the peripheral positions of the stimuli to which they responded during the experiments. The range of mean reaction-times under normal conditions was from 130 to 250 sigma. Large individual differences between subjects were found in all the studies, but these were lost sight of as a result of pooling the data. Reaction-time of the eye has been shown to change with repetition. In general it becomes shorter and this effect persists over a relatively long period of time. Whether or not the shortening in latency is a genuine practice effect is open to question. Dodge suggested that the arrangement of stimuli (number of positions used and their location in the visual field) might be partly responsible for this effect.

Dodge and Miles later incorporated the measurement of ocular latency into various studies of the effects of restricted diet and alcohol on human behavior. Reduced diet had no effect on eye latency. In regard to moderate doses of alcohol, the general conclusion is that the effect is a relative lengthening of latency amounting to approximately five percent. The present writer would like to suggest that more information is

[1] This study was carried out at the Psychological Laboratories of the University of Minnesota.

[2] W. R. Miles. The reaction-time of the eye. *Psychol. Monog.*, 1936, **47**, 268–293.

needed concerning normal ocular reaction-time and factors associated with its variability if it is to be used to study the effects of alcohol, low diet, drugs, anoxemia, etc. We now have available much more adequate experimental designs to study the significance of observed effects of such factors on behavior.

Miles investigated the role of the stimulus situation in relation to changes in ocular latency. Although his conclusions were tentative, it appears that size of the stimulus, variation in stimulus position, distance of the peripheral stimulus from the preëxposure fixation point, and monocular versus binocular conditions of vision are associated with changes in latency. Further work on the role of such stimulus variations as determinants of mean latency and variability is needed.

The purpose of the present study is to investigate the latency of ocular response to suddenly appearing peripheral stimuli when the line of regard is shifted to a new stimulus position. The following factors were chosen for investigation: variation in the zonal position of the stimulus, repetition, knowledge or lack of knowledge of stimulus position, and individual differences in response.

This study represents a departure from the usual experimental design and statistical treatment employed in studies of reaction-time. Previous investigations of ocular latency have led to rather equivocal results regarding the effects of repetition, individual differences, and zonal position of the stimulus on reaction-time. The single variable technique appears to be an inadequate and inefficient method of design which does not lead to effective use of all the information latent in the data. The question at hand is, to what extent do these factors enter into the observed variability in the reaction-time of the eye? Our hypothesis is that all of the observed reaction-times are drawn from a homogeneous population (null hypothesis). If this hypothesis is not refuted, we may interpret our results to mean that stimulus position, repetition, and individual differences do not affect ocular latency. We wish to make a valid estimate of *experimental error*,[3] in order to

[3] Residual variation about the mean which has been freed from variation traceable to ascribed sources, *e.g.* position in this case.

make a valid test of the variation among position means, *e.g.* for zonal positions. Too often estimates of the experimental error do not represent residual variation which is random, since no attempt is made to free it from that part of the variability whose source is known.

Given the proper experimental design, the *analysis of variance*[4] can be used to segregate from comparable groups of data the variation traceable to specified sources. By means of the associated tests of significance, a judgment can be made as to whether or not several groups are samples from a single, homogeneous population. In the present study, for example, variation can be traced to the factors mentioned above. The obtained experimental error is freed from that part of the variability which has been traced to those sources. It represents residual random variation plus variation due to factors which are unknown or uncontrolled under the given experimental conditions. By means of the experimental error we can determine the probability that the variation among position means, for example, is greater than that within the whole group of data. Further discussion will be postponed until the presentation of the results.

Apparatus and Method

The Minnesota eye-movement camera was used for photographic recording of eye-movements. Both observation and recording were binocular. A time record was obtained by interrupting the recording beam 60 times per second with a blade mounted on a synchronous motor. The beginning of the reaction-time was indicated by means of a solenoid arrangement attached to a vane. At the 'ready' signal the vane was depressed manually, cutting off the recording beam. At the exact instant the stimulus light came on, the vane was released by means of a double-pole, double-throw switch, allowing the camera to begin recording the latent period. The end of the reaction-time was marked by the beginning of the eye-movement in response to the stimulus.

The stimuli used in this study were lights in a dark field. They (ordinary flash-light bulbs) were mounted separately in a metal box placed 33 cm in front of the subject. The face of the box had one hole in its center and eight holes arranged symmetrically around it on a circle whose radius was 9.5 cm. Each hole (6 mm in diameter) had a light placed in back of it. The visual angle from the center fixation to a peripheral light was approximately 16 degrees. Hereafter these positions will be

[4] R. A. Fisher. *Statistical Methods for Research Workers.* London: Oliver and Boyd, 1936. Pp. xiii, 339. G. W. Snedecor. *Calculation and Interpretation of Analysis of Variance and Covariance.* Ames: Collegiate Press, 1934. Pp. 96. *Statistica Methods.* Ames: Collegiate Press, 1937. Pp. xiii + 341.

referred to by number as follows: No. I is up; No. III is upper right; No. V is right; No. VII is lower right; No. IX is down; No. XI is lower left; No. XIII is left; No. XV is upper left. A unique feature of the set-up was that the preëxposure fixation light (center light) was automatically turned out when the stimulus light flashed. Thus there was never more than one light in an otherwise dark field to compete for attention.

Rather than use a large group of subjects with a few reactions from each, a group composed of eight college women and eight college men was used, with a large number of reactions from each subject. These students were drawn at random from sophomore classes in laboratory psychology.

The subjects were given the following instructions:

> "Fixate the center dot. After a brief interval the center light will disappear and a new light will appear somewhere near the outside of the reading table. As soon as this new light appears you are to fixate it as quickly and as accurately as possible. Keep looking at it until it disappears. That is, keep your fixation accurately on the center light until the new one appears, and then shift to the new light as quickly as possible. Be sure that you look directly at each spot of light as long as it is on. As soon as the outside light disappears you are to shift back to the center light and be ready for the next movement."

Each subject gave two sets of 80 reactions, a set being arranged in ten series of eight reactions each. Each series had in it one each of the eight different stimulus positions in a random order from one series to the next. This permits the segregation of the effects of repetition and position separately in the analysis of variance. It also allows a study of the effect of repetition, with the effect of position randomized.

Another feature of the experiment was the division of the total number of observations into two sets. In one, the subject did not know the zonal position of the stimulus—this will be called 'without knowledge.' In the other, the subject knew the zonal position at which the stimulus was to appear. The order of the stimuli was the same in the two situations. Part of the subjects were given the 'without knowledge' on one session and the 'with knowledge' the second session. This gave a counterbalanced order intended to neutralize the practice effect or a possible interference effect from one to the other.

Results: Mean Latencies

It seems appropriate to present the results of this study as shown by customary methods of analysis before presenting those from the analysis of variance, since the former are more easily compared with those previously obtained by other experimenters.

Knowledge of Stimulus Position

Since each subject observed under two different conditions, it is possible to make a direct comparison of the reaction-times when : (1) S knew the location of the next peripheral stimulus, and (2) when S did *not* know the location of the next stimulus. It will be recalled that half of the S's observed

under the 'without knowledge' conditions in the first sitting and 'with knowledge' in the second sitting; the reverse being true for the other half. This was done in the expectation of cancelling out an overall practice effect from the first day to the second.

The results are summarized in Table I. All records were read in 1/120 sec., the error of observation in reading the records being not over 1 unit. Raw scores and machine methods of computation were used throughout.

TABLE I

COMPARISON OF MEAN LATENCIES UNDER VARIOUS EXPERIMENTAL CONDITIONS
(*In 1/120 Sec.*)

Reaction	*N*	Mean	Diff.	C.R.
With Knowledge—First Day	491	22.58	3.20	6.74
With Knowledge—Second Day	632	19.38		
Without Knowledge—First Day	596	24.53	.30	—
Without Knowledge—Second Day	520	24.23		
Total: With Knowledge	1123	20.78	3.62	11.00
Total: Without Knowledge	1116	24.40		

There appears to be a highly significant difference between reaction-times obtained under the two experimental conditions. The mean value for reaction-time *without knowledge* is about 200 sigma, a finding which agrees with previous studies. The mean value *with knowledge* is about 165 sigma. However there is a great deal of overlapping. There are more abnormally long reactions in the former and many more very fast reactions in the latter. The test of association (chi square) shows that these two distributions are not samples from the same population, *i.e.*, there is a significant difference between groups. Comparison of the variances by Snedecor's *F* test shows that the variability is significantly greater when the Ss observed with knowledge of the next stimulus position.

Detailed analysis of the results brings to light facts which lead us to suspect that there are several factors which are contributing to variability in the observed reaction-times and obscuring the obtained results. Probably the most important

is that there are marked individual differences among the subjects when reaction-times are compared under the two experimental conditions. Since the results are not consistent from subject to subject, this appears to be an important source of variation in reaction-time which must be taken into account. An interesting finding is that the effect of practice (so-called) from the first experimental day to the second is not the same for the two types of observation. Subjects who observed with knowledge on the second day show a significantly faster mean reaction-time than those observing under the same conditions on the first day. However, it appears not to make any difference in mean latency when the subject is without knowledge of the next stimulus position; the mean reaction-times are the same, whether on the first sitting or the second. Thus, familiarity with the experimental conditions and practice in moving the eyes as rapidly as possible may lead to a reduction in latency, only under certain conditions. A third finding which has some bearing on the point in question is that all of the stimulus positions save one (No. V, right) show significantly faster reaction-times for the *with knowledge* conditions. This particular response (No. V) is a highly practiced one, since it is the same as the saccadic movements in reading.

The Effect of Repetition

Previous studies have been inconclusive concerning the effect of practice, although one cannot deny that eye-latency does (on the average) become shorter with repetition. In this study a comparatively large number of reactions from each observer made it possible to study changes in reaction-time with repetition. Previous studies have used a comparison of first-half against last-half to test for practice effect. We have analyzed the results for individual subjects, using Fisher's t for small samples. The following results were obtained: The general trend is in the direction of a lowering of the latency, but individual differences are very marked. It seems likely that whether or not the practice effect occurs (as judged by this criterion) depends to a large extent on the individual subjects.

Since the original observations were arranged in series with all eight positions repeated once in each, it is possible to study the practice effect in other ways. A mean value was calculated for each series under both experimental conditions. These values were used to draw practice curves, which are shown in Fig. 1. In each of the four sets of means there is a

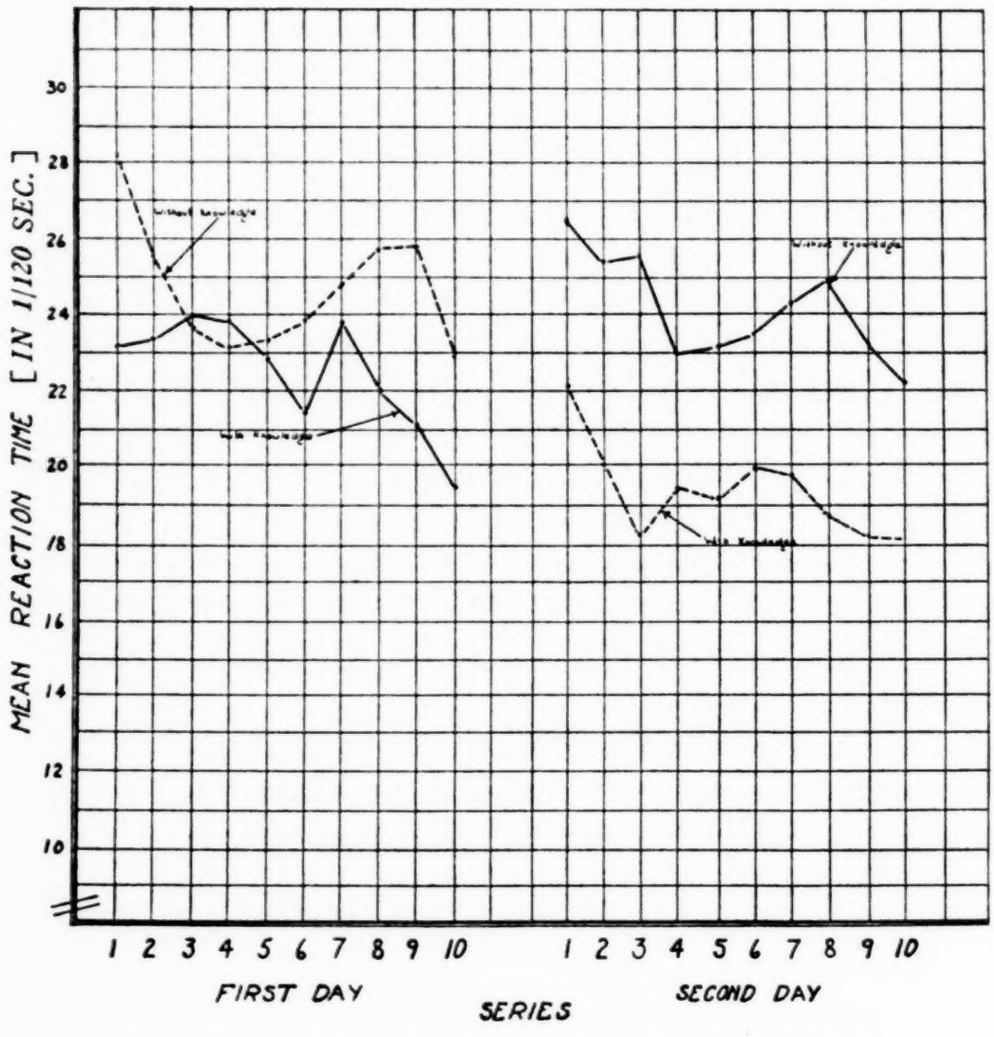


FIG. 1. Effect of practice on latency of the eye [pooled data].

decrease from Series 1 to Series 10, showing a general trend as regards the effect of repetition on reaction-time of the eye. If we compare the first and last series for each separate curve, we find that in only one case is the difference highly significant. That is for without knowledge, first experimental day. In the other three cases, the differences are in the expected direc-

tion, but further data are needed before definite conclusions can be drawn. Practice curves drawn for the individual subjects show the same trend as the pooled data, but are by no means consistent. Only two of the subjects show a significant practice effect for both with-and-without knowledge. Whether or not there is a significant decrease in latency with repetition seems to depend on the particular subjects being considered.

Zonal Position of the Peripheral Stimulus

The data for position were pooled and measures of central tendency and variability computed for each of the eight

TABLE II

Summary Table Showing the Effect of Position on Reaction-Time of the Eye

(*In 1/120 sec.*)

Position	First Day With Knowledge			Second Day With Knowledge		
	N	Mean	S.D.	*N*	Mean	S.D.
I	57	23.30	9.85	63	17.59	5.53
III	63	22.51	7.63	80	17.43	6.75
V	62	23.73	9.22	82	19.44	6.72
VII	62	21.65	6.65	83	20.83	8.58
IX	62	22.69	6.18	81	21.32	8.37
XI	62	22.63	7.77	84	19.15	7.47
XIII	62	23.26	8.73	81	20.99	9.05
XV	61	20.95	7.14	78	17.79	7.34
Position	**First Day Without Knowledge**			**Second Day Without Knowledge**		
	N	Mean	S.D.	*N*	Mean	S.D.
I	73	23.18	6.46	63	24.94	8.55
III	74	21.96	5.55	68	24.00	7.75
V	76	23.11	7.24	65	23.49	6.41
VII	73	26.01	8.86	67	22.97	6.16
IX	76	26.80	7.12	65	24.00	6.99
XI	75	26.00	7.72	63	24.27	6.48
XIII	74	26.77	9.05	63	25.86	6.83
XV	75	22.47	6.58	65	25.03	9.20

stimulus positions. These are shown in Table II. Examination of this table reveals differences between the mean reaction-time for the various stimulus positions. It is apparent,

however, that the least variation between position means occurs when O observes with knowledge. The largest variation between position means is for the observations without knowledge on the first experimental day. Grouping all positions into upper field vs. lower field shows significant differences in favor of the upper field for both *without knowledge* and *with knowledge* reactions. Other group comparisons are shown in Table III. None of the other differences are highly significant.

TABLE III

COMPARISONS OF VARIOUS PORTIONS OF THE VISUAL FIELD

Comparison	Reaction	Field	*N*	Mean	Diff.	C.R.
Upper Field * vs. Lower Field	Without Knowledge	Upper	418	23.52		
		Lower	419	25.08	1.56	3.02
	With Knowledge	Upper	402	19.69		
		Lower	434	21.24	1.55	2.88
Right Field † vs. Left Field	Without Knowledge	Right	423	23.58		
		Left	415	25.06	1.48	2.83
	With Knowledge	Right	432	20.71		
		Left	428	20.61	.10	.18

* Nos. V and XIII omitted.

† Nos. I and XI omitted.

It would seem reasonable to expect differences between position means in both types of reactions if the variation in reaction-time according to position is referable to some physiological factor, such as differential retinal sensitivity. However, if it is a result of habit or learning, it may be present to a greater extent in the reactions where S does not know in which direction the next stimulus will come. The latter appears to be the case, but before we can draw any conclusions it is necessary to have an adequate test of the significance of the variation among the position means. If position is not associated with reaction-time, the means for the positions should vary only within the limits of random sampling. Whether or not the null hypothesis (ibid.) is refuted in this case is a question to be answered by the analysis of variance.

The Analysis of Variance

We have presented an analysis of the results for reaction-time using the customary methods of statistical treatment. In the case of repetition, the results are equivocal. In the case of position, differences were found between the different stimulus locations. Further, individual differences among subjects have been shown to enter into and obscure the results. The question arises as to what extent these three factors account for the observed variability in reaction-time of the eye. Up to this point we have not made use of all the information latent in the data, nor have we obtained an answer to the question which the experimental design affords. The hypothesis in this type of experiment is that all of the observed latencies are drawn from a homogeneous population. If this hypothesis is not refuted, it follows that position, repetition and individual differences do not affect reaction-time. In addition, we need a valid estimate of experimental error in order to make our tests of significance for the variation among sample means for positions, etc. Since the four experimental conditions (with knowledge and without knowledge on either the first or second day) have been found to exhibit rather marked differences, separate analyses have been made for each one.

TABLE IV

Analysis of Variance without Knowledge—First Day

Source of Variation	df*	Sum of Squares	Mean Square	*F*
Total	595	34,848.1130	58.57	
Between Means of Series	9	1,229.9690	136.66	3.79†
Between Means of Subjects	8	10,834.5499	1,354.32	37.54†
Between Means of Positions	7	2,182.6045	311.80	8.64†
Remainder	571	20,600.9896	36.08	

* Number of degrees of freedom.

† Highly significant ($P < .01$).

Table IV shows the analysis of variance for without knowledge, first day. We find that the total variance has been reduced from 58.57 to 36.08, when that part of the variability whose source is known has been removed (*i.e.*, between means

of series, subjects and positions). The latter value, called the experimental error, is used to test the significance of the variation among group means. It represents the residual variation within groups in the whole set of data. When Snedecor's F test is used,[5] we find that the variation between means of positions, series, and subjects is highly significant in each case. The largest source of variation present is between means of subjects, *i.e.*, individual differences among the subjects. About one-third of the sum of squares is associated with this source of variation.[6] Since the variation between means of positions is highly significant, we conclude that zonal position of the peripheral stimulus has a significant effect on the reaction-time of the eye when S does not know the direction in which the new stimulus will be exposed.

Comparable results for the other three experimental conditions (plus those in Table IV for comparison) are shown in Table V. In every case the variation between means of subjects is highly significant. It is the major source of variation in reaction-time of the eye in this study. As a matter of fact the efficiency of the experiment is seriously impaired when this source of variation is neglected. When S observed with knowledge on the first experimental day, the variation between means for series is not significant (indicating no practice effect), nor is that for variation between means of positions significant. The experimental error is reduced from 64.31 to 58.38, indicating that the experimental design failed to take into account other important sources of variability. When S observed with knowledge on the second experimental day, the variation between means of positions is highly significant, but that between means of series is open to suspicion (P lies between .05 and .01). Results for observations without

[5] F is the ratio of the larger variance to the smaller variance; *e.g.*, 311.80/36.08. The mean square for the remainder (error) is used to test the significance of the variation of each group of means. P is the probability of getting as large an F value by errors of random sampling. Thus when $P<.01$ we know that the variation between means for position, for example, is highly significant—*i.e.*, reaction-time of the eye varies with different stimulus positions.

[6] The term 'analysis of variance' is really a misnomer, since it is the 'sum of squares' which is divided up into its constituent parts together with the number of degrees of freedom.

TABLE V

SUMMARY OF INFORMATION FROM ANALYSES OF VARIANCE FOR REACTION-TIME

Source of Variation	With Knowledge				Without Knowledge			
	First Day		Second Day		First Day		Second Day	
	df	Mean Square	df	Mean Square	df	Mean Square	df	Mean Square
Between Means of Positions....	7	50.95†	7	199.73*	7	311.80*	7	222.48*
Between Means of Series.....	9	108.63†	9	85.32‡	9	136.66*	9	91.95*
Between Means of Subjects....	6	475.49*	8	1704.53*	8	1354.32*	6	1844.35*
Remainder (Error).....	468	58.38	607	36.71	571	36.08	497	32.97

* Highly significant ($P < .01$).

† Not significant ($P > .05$).

‡ Open to suspicion ($P < .05 > .01$).

knowledge on both first and second days reveal highly significant differences between all means compared.

It seems reasonable to assume that if the position effect depends on some purely physiological factor it ought to be operative in all of the experimental conditions. Thus, whether or not the subject knows where the light is coming next should not make any difference. Our findings, however, lead to the hypothesis that the position effect depends on such psychological factors as learning or previous experience, rather than differential sensitivity. The results lead to the further hypothesis that the sources of variation in reaction-time of the eye are psychological rather than physiological. The determination of the nature and effects of these factors remains a subject for future experimentation, particularly as regards reactions where S has knowledge of the stimulus conditions.

CONCLUSIONS: ANALYSIS OF VARIANCE

1. The variation among means of subjects is highly significant for all experimental conditions and is a major source of variability in the observed reaction-times.

2. Reaction-time is affected by variation in zonal position

of the peripheral stimulus, except in the case where a naive subject knows the position of the stimulus before he reacts to it.

3. Reaction-time of the eye varies, depending on the amount of practice the subject has had, when he has no exact knowledge of the position of the stimulus to be reacted to. This result is for pooled data and is not consistent from subject to subject, as pointed out previously. The general trend is for latency to decrease with practice. In the case of without knowledge, first day, the apparent drop in latency as shown in Fig. 1 is misleading since the variation between means for the series is not significantly greater than that within the series.

(Manuscript received June 4, 1940)

THE EFFECT OF SEQUENCE OF CONTINUOUS AND PERIODIC REINFORCEMENT UPON THE 'REFLEX RESERVE'

BY FRED S. KELLER

Department of Psychology, Columbia University

Extinction curves of the lever-pressing response in rats have been shown by Skinner (**5**) to possess properties that depend in part upon the mode and amount of preceding reinforcement. When each of, say, 100 responses made by a given animal has been reinforced with food, a subsequent curve of extinction (*a*) has a high initial acceleration and a later rapid retardation of response rate; (*b*) shows cyclic 'emotional' deviations, due to recurring intervals of response decrement and 'compensatory' increment; and (*c*) contains a total response number (following 100 reinforcements) that probably does not exceed 200.

When, however, reinforcement of the lever depressions has been *periodic*, rather than regular or continuous, the extinction curve shows markedly different characteristics. If a rat's responses to the lever have been reinforced at 5-minute intervals only, (*a*) neither the initial acceleration nor the ensuing decline in response rate during extinction is as great as it would be if all the preceding responses had been reinforced; (*b*) the cyclic deviations in the curve are absent, presumably because of emotional adaptation during the periodic reinforcement; and (*c*) there is a tremendous increase in the number of responses that appear before extinction is at all complete (500 responses from a subject are not unusual after prolonged periodic reinforcement).

These and other experimental operations have led Skinner to the concept of 'reflex reserve,' which is treated as a reserve of responses implanted within the organism by the process of reinforcement—a reserve the properties of which are indicated only during its exhaustion, when reinforcement is withdrawn

and extinction takes place. In the cases mentioned above, the two types of extinction curve show the mode of reinforcement to be an important determiner of the *size* of this reserve as well as the *proportionality* between the reserve (the potential responses) and the 'reflex strength'—the rate of responding at any chosen stage of extinction.

It has also been shown by Skinner that when periodic reinforcement *follows* continuous reinforcement prior to extinction it remains efficacious in adding to the reserve and changing the proportionality between the reserve and the strength. This is, perhaps, to be expected from the nature of the operations involved, since periodic reinforcement demands at least a single initial reinforcement which, if repeated regularly for each lever-pressing response, would constitute a preliminary period of continuous reinforcement. In other words, the difference is basically between one and more-than-one initial reinforcements prior to the introduction of intervals in which no responses are reinforced.

The effect of the alternative sequence of operations—periodic *followed by* continuous reinforcement—has not, however, been investigated, and several consequences of such a procedure suggest themselves. There is the possibility (1) that neither the size of the reserve nor the proportionality between reserve and strength will be affected by the sequence of continuous and periodic reinforcement—that extinction curves for animals subjected to the continuous-periodic sequence will be essentially the same as those for animals subjected to the periodic-continuous sequence; (2) that both the size of the reserve and the proportionality will be affected; (3) that the size of the reserve will be affected, but the proportionality will be unchanged; or (4) that the size of the reserve will be unaffected, but the proportionality changed.

The realization of one of these possibilities, which is the concern of the present study, should be of some significance with respect to the concepts of reserve and strength. There is nothing in Skinner's formulation which would lead one to expect that the *size* of the reserve is dependent upon the sequence of reinforcement as outlined above, but there are not,

on the other hand, any data which would indicate that such is *not* the case, and a discovery of such a relation would have rather important theoretical and practical implications.

Twenty male white rats, 75 to 100 days old, were conditioned to depress a lever for small pellets of food, uniform in size and composition, in a mechanically-operated form of the Skinner apparatus (see 4 and 5). In the training procedure each rat was introduced to the response chamber, following a food-deprivation of 22 hours, and permitted to eat one or two grams of food already placed in the food-tray by the experimenter. During this time the lever above the tray was not yet connected with the food-magazine in the adjoining compartment. After the first food supply was exhausted, the experimenter caused several pellets to be discharged, one at a time, from the magazine into the tray. When the stage was reached at which these pellets were eaten as soon as they appeared, the magazine was connected to provide reinforcement for each lever-pressing response made by the animal.

The criterion of conditioning was considered to be satisfied by a given rat when his rate of lever-pressing became relatively constant and rapid. Since at least 15 responses were to be reinforced *after* conditioning, the determination of a 'learning-point' was not as difficult as it might appear, although some lack of precision was unavoidable. Stop-watch measurement of the times between successive responses, together with inspection of the cumulative response-frequency curve made by the animal (and recorded on a slow-moving kymograph drum outside the apparatus), provided a sufficiently accurate designation of the point at which conditioning had been established and permitted the experimenter to calculate the number of remaining reinforcements to be given.

The animals were advanced from one stage of training to another as fast as their behavior seemed to warrant, with the result that some required fewer daily sessions than others before the habit was formed. Nine rats were conditioned on the first day, in less than an hour; eight on the second day; two on the third; and one on the fourth. In the case of four rats 40 successive responses were reinforced immediately after the satisfaction of the criterion; the remainder were given only 15 reinforcements directly after mastery, but received an additional 25 on the following day. Thus, each animal received a total of 40 reinforcements beyond those required for conditioning.

Subsequent to this preliminary training the animals were divided into two groups of 10 each, equated as well as possible on the basis of litter membership, body weight, and time required for learning. The rats of Group A were first given two days of periodic reinforcement at 3-minute intervals. This was followed by two more daily sessions in each of which 30 successive responses were reinforced. Finally, on the next day, a 3-hour extinction curve was obtained from each animal.

The sequence of periodic and continuous reinforcement was reversed for the rats of Group B. Two days of *regular* were followed by two days of *periodic* reinforcement, with a final day in which the response was extinguished. In all other essential respects the conditions for the two groups were identical. The experimental sessions took place daily, following 22-hour deprivation; the general laboratory environment was fairly constant with respect to light, noise, temperature, and other factors; and the animals of the two groups were matched with respect to response-compartments (three Skinner boxes were used during the experiment). Such conditions should be adequate for the maintenance of drive constancy and the avoidance of variations in the records that might be traced to 'situational' sources (4).

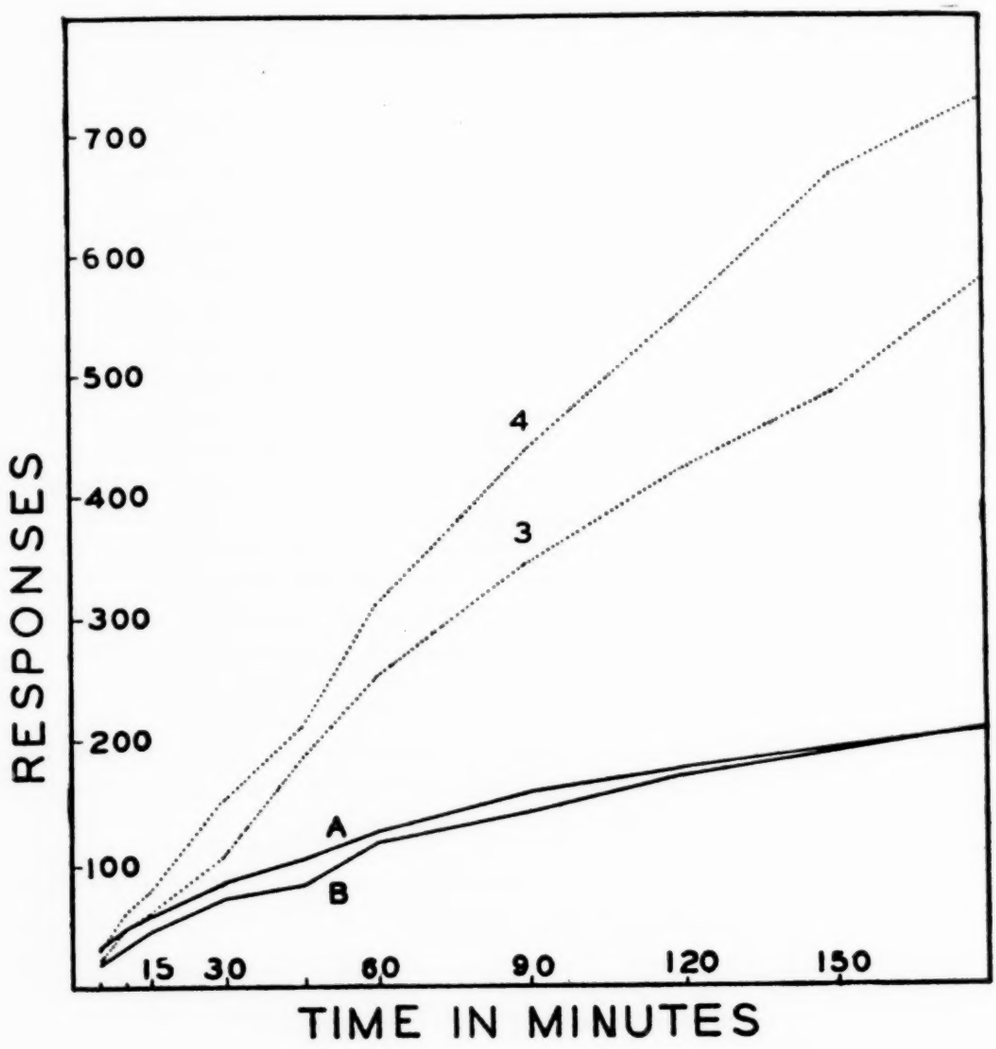


Figure 1 presents composite extinction curves for the two groups of animals. These curves are based on the data for 16 rats. The records of two animals of each original group were not comparable to those of the other subjects beyond the first 15 minutes of responding, since their extinction was not continuous during the 3-hour period. Inclusion of these data, which were obtained in alternate 15-minute periods, does not appear to be justified.

The *t*-test was used to compare the average number of responses of the two groups throughout the extinction period. Table I is a summary of this comparison.

The *t*-values necessary for 5.0 percent and 1.0 percent significance, with 14 degrees of freedom, are 2.14 and 2.98 respectively (3). Table I indicates that Group A, during the

TABLE I

SHOWING THE MEAN NUMBER OF RESPONSES OF GROUPS A AND B AT DIFFERENT STAGES OF EXTINCTION, THE DIFFERENCES BETWEEN THE MEANS, AND THE CORRESPONDING *t*-VALUES

Time, in Minutes, from Beginning of Extinction . . .	5	10	15	30	45	60	90	120	150	180
Mean responses, Group A	30.7	45.0	56.1	84.9	104.6	124.9	156.2	177.6	194.1	208.4
Mean responses, Group B	17.2	33.0	43.9	70.4	99.1	118.2	141.7	174.7	192.9	209.7
Difference between means	13.5	12.0	12.2	14.5	5.5	6.7	14.5	2.9	1.2	1.3
t-values	3.34	1.79	1.66	1.22	.414	.438	.729	.122	.047	.045

first five minutes of extinction, responded a significantly greater number of times than Group B, since the *t*-value is greater than that required to reach the 5.0 or 1.0 percent point. Thereafter, the difference between groups is not significant.

This is at least a rough indication that the effect upon the curves of Fig. 1 is felt primarily at their beginning, during the first few minutes of the records. Because of the small number of subjects and the arbitrary termination of the extinction at the end of three hours, a more detailed examination of the data is inadvisable.

The bearing of these results upon the question raised at the beginning of this report is clear. Of the suggested possibilities, the fourth has been realized. The effect of reversing the sequence of continuous and periodic reinforcement, under the conditions of this experiment, is (*a*) insignificant with respect to the *size* of the reflex reserve, but (*b*) shows itself as a changed proportionality between the reserve and the reflex strength. In other words, the total number of available responses is neither increased nor decreased as a result of the sequence chosen, but the rate of emission is altered.

In his description of the effect of periodic, as against regular or continuous, reinforcement, Skinner notes that "we

cannot compare the heights of extinction curves (at some arbitrary point) without taking the preceding reinforcement into account." This statement is amply substantiated in the present study, and a degree of elaboration is justified: not only must we consider gross differences in mode of reinforcement, but also differences with respect to temporal arrangement of these modes when both are applied prior to extinction.

Although the above interpretation of the results obtained in this experiment with the two main groups of animals is entirely consistent with Skinner's concepts of 'reserve' and 'strength,' some reference should be made to the anomalous behavior of those rats the records of which have been omitted from the composite curves. Extinction for rats 1, 2, 3, and 4 was discontinuous, taking place in 15-minute sessions at 15-minute intervals. The records for rats 1 and 2 are incomplete and may be neglected here, but 3-hour totals for rats 3 and 4 (of groups B and A respectively) are available and are represented by dotted lines in Fig. 1. Whereas the average number of responses during extinction for the 16 rats of groups A and B was 209 (range, 91–354), the response-totals for rats 3 and 4 respectively were 576 and 723. Moreover, the shape of the cumulative curves, under such 'spaced' extinction, is strikingly different from that which characterizes the 'normal' or continuous curve, either as obtained by Skinner or for groups A and B of the present study.

Comparable effects of the discontinuous extinction procedure have apparently been observed by Ellson (**2**), who has used his findings as the basis of an attack upon the 'reflex reserve' concept and, by extension, the entire descriptive system proposed by Skinner. The need for adequate analysis of such a seeming discrepancy is obvious, but will not be undertaken here. Experimental data which bear upon this question will, however, be presented in a later paper.

A conceivable effect of the sequence-reversal, although one not demonstrated in this experiment, may be mentioned briefly. In view of the known difference in size of reserve created by the two types of reinforcement (periodic and continuous), one might entertain the view, prior to actual in-

vestigation, that the periodic-continuous sequence would operate to *reduce* the number of responses during extinction. Such an effect, if obtained, would bear a specious resemblance to Dunlap's findings with respect to the breaking of undesirable habits (1) and, at the same time, be consistent with the view that the rat stops responding sooner because he has come to expect a reward for *every* lever-depression. Although it is unlikely that the results herein reported will escape the generality of 'expectancy' interpretation, it is quite certain that they do *not* provide us with a new technique for the quicker removal of well-entrenched habits—for subtracting from the reflex reserve by the simple operation of reinforcement.

(Manuscript received May 29, 1940)

References

1. Dunlap, K., *Habits: their making and unmaking*, N. Y., Liveright, 1932, x + 326.
2. Ellson, D. G., The concept of reflex reserve, *Psychol. Rev.*, 1939, **46**, 566–575.
3. Goulden, C. H., *Methods of statistical analysis*, N. Y., Jno. Wiley & Sons, 1939, p. 267.
4. Keller, F. S., Some experimental operations affecting drive constancy, *J. comp. Psychol.*, 1939, **28**, 63–72.
5. Skinner, B. F., *The behavior of organisms*, N. Y., Appleton-Century, 1938, pp. 457.

SPEED AS A DETERMINER OF MUSICAL MOOD

BY MELVIN G. RIGG

Oklahoma A. and M. College

From previous experiments it has appeared that tempo has more influence upon the emotional suggestiveness of music than any other single factor.[1] This feature would therefore seem to deserve a more exhaustive study than it has hitherto received. The present paper reports an experiment dealing with speed alone.

The stimuli consisted of five musical phrases played on the piano. The phrases were harmonized and were four measures in length.[2] Each was presented at six different speeds, the tempo being rigidly controlled by a metronome.[3] The phrases are designated as A, B, C, D, and E with a number following to indicate the speed (the number of quarter notes to the minute).

The order of playing was as follows:

(read down)		
E 120	E 138	E 160
D 120	D 138	D 152[4]
C 120	C 138	C 160
B 120	B 138	B 160
A 120	A 138	A 160
E 100	E 80	E 60
D 100	D 80	D 60
C 100	C 80	C 60
B 100	B 80	B 60
A 100	A 80	A 60

The observers in the experiment were 88 college students enrolled in psychology classes. They represented varying de-

[1] Rigg, M. G., What features of a musical phrase have emotional suggestiveness? *Pub. Soc. Sci. Res. Coun. Oklahoma A. and M. Coll.*, 1939, 1, 38 pp.

[2] These phrases were used in the study previously mentioned, *op. cit.* They may be found in musical notation in the appendix, designated as A 1, B 1, C 1, D 1, and E 1.

[3] It was at first feared that the sound of the metronome would have an influence upon the emotional effect of the music, but preliminary experimentation revealed that such was not the case.

[4] It was beyond the technical ability of the pianist always to give a perfect rendition of this phrase at 160.

grees of musical training, but it seems unlikely that musical training influences judgments concerned with the emotional effects of music.[5]

Each auditor was supplied with a check list to assist him in making his replies. From previous studies, the most valid discrimination with regard to the emotional character of music appeared to be between the happy and the sad. The observers were consequently asked to designate whether the phrase was primarily sad-serious (indicated as *S*) or pleasant-happy (indicated as *Y*). After this judgment the auditors were to attempt a finer discrimination by classifying the phrase under subheadings as follows:

X Serious-sad
- 1 Solemnity
- 2 Sorrowful longing
- 3 Melancholy
- 4 Lamentation
- 5 Agitation

Y Pleasant-happy
- 6 Hopeful longing
- 7 Love
- 8 Revery
- 9 Gaiety
- 10 Joy
- 11 Triumph [6]

It was explained to the observers that melancholy was to mean something less intensely sad than lamentation, and gaiety, something less intense than joy.

The data for Phrase A are presented in Table I. The results for the various renditions of this phrase are given not in the order of performance, but in that of increasing speed. When this phrase was played 60 quarter notes to the minute, the *X* votes were unanimous, and 37 of these were specifically

[5] Hevner, K., The affective character of the major and minor modes in music, *Amer. J. Psychol.*, 1935, 47, 103–118. Rigg, M. G., An experiment to determine how accurately college students can interpret the intended meanings of musical compositions, *J. exper. Psychol.*, 1937, 21, 223–229.

[6] This list is identical with the one used in the previous investigation. See first footnote.

TABLE I

PHRASE A

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	88	86	86	86	80	75
1. Solemnity	24	22	20	22	11	12
2. Sorrowful Longing	7	7	9	10	4	3
3. Melancholy	15	14	12	23	17	8
4. Lamentation	37	34	32	19	21	11
5. Agitation	5	9	13	12	27	41
Y Pleasant-Happy	0	1	2	2	8	13
6. Hopeful Longing	0	0	0	2	5	7
7. Love	0	1	0	0	0	0
8. Revery	0	0	0	0	1	4
9. Gaiety	0	0	2	0	1	0
10. Joy	0	0	0	0	0	0
11. Triumph	0	0	0	0	1	2

for lamentation. The *X* responses remain at 86 for the three next faster speeds, but the lamentation votes go down successively to 34, 32, and 19. It is safe to conclude therefore that the phrase is progressively less intense in its sadness. At the two faster speeds (138 and 160) the *X* responses go down, first to 80 and then to 75. Even at the fastest speed the phrase is still sad, but the judgment is no longer unanimous and there is a tendency for the music to suggest agitation.

The data for Phrase B are presented in Table II. This

TABLE II

PHRASE B[7]

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	32	24	15	9	6	5
1. Solemnity	3	2	1	0	0	0
2. Sorrowful Longing	5	7	3	1	1	0
3. Melancholy	3	1	2	0	0	1
4. Lamentation	6	3	1	2	1	1
5. Agitation	15	11	8	6	4	3
Y Pleasant-Happy	54	63	73	79	82	83
6. Hopeful Longing	23	19	9	4	2	2
7. Love	2	7	5	3	1	3
8. Revery	12	13	11	6	7	10
9. Gaiety	14	17	31	40	31	20
10. Joy	2	5	13	22	34	42
11. Triumph	1	2	4	4	5	6

[7] Since occasionally an auditor would mark a phrase *X* or *Y* without making the second discrimination, the total for the subheadings may be slightly less than the number for the main heading.

phrase shows a different picture, since it is primarily pleasant-happy in character. For every increase of speed the *Y* votes become more numerous. Between the two fastest speeds this gain is only 1; nevertheless, the joy responses advance from 34 to 42.

TABLE III

PHRASE C

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	81	81	79	67	58	43
1. Solemnity	6	10	7	11	10	6
2. Sorrowful Longing	31	26	25	24	17	11
3. Melancholy	26	30	29	19	12	4
4. Lamentation	15	15	17	11	8	6
5. Agitation	3	0	1	2	11	16
Y Pleasant-Happy	7	7	9	21	30	45
6. Hopeful Longing	4	5	2	9	13	22
7. Love	1	2	4	5	7	7
8. Revery	2	0	3	4	7	9
9. Gaiety	0	0	0	2	2	4
10. Joy	0	0	0	1	0	2
11. Triumph	0	0	0	0	1	1

The data for Phrase C are presented in Table III. This phrase is definitely sad except at the fastest tempo. The changes between the two slowest speeds appear inconsequential, but the *Y* votes show a steady gain at the fastest speeds.

The data for Phrase D are presented in Table IV. At slow speed this phrase is sad; the faster it is played, the happier it becomes. At the fastest tempo it suggests triumph.

TABLE IV

PHRASE D^7

Metronome Speeds	60	80	100	120	138	152
X Serious-Sad	50	44	19	13	4	1
1. Solemnity	13	16	8	6	1	0
2. Sorrowful Longing	22	15	7	4	0	0
3. Melancholy	4	5	2	2	1	0
4. Lamentation	8	6	1	1	1	0
5. Agitation	2	2	1	0	1	1
Y Pleasant-Happy	38	44	68	75	84	87
6. Hopeful Longing	18	15	27	17	8	4
7. Love	1	6	5	2	2	2
8. Revery	4	7	11	15	7	4
9. Gaiety	0	1	3	11	5	7
10. Joy	2	3	4	4	17	11
11. Triumph	13	12	17	26	45	59

The data for Phrase E are presented in Table V. Here there is one inversion, since for speed 160 there are only 85 *Y* responses as against 87 for speed 138. However, gains are registered in both the categories of joy and gaiety for the faster of these tempos.

TABLE V

PHRASE E

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	41	33	19	5	1	3
1. Solemnity	11	9	5	2	0	0
2. Sorrowful Longing	16	15	9	2	0	1
3. Melancholy	8	6	4	0	0	1
4. Lamentation	4	3	1	1	1	0
5. Agitation	2	0	0	0	0	1
Y Pleasant-Happy	47	55	69	83	87	85
6. Hopeful Longing	16	27	30	29	14	11
7. Love	18	15	26	18	16	4
8. Revery	13	13	11	28	18	8
9. Gaiety	0	0	1	5	25	33
10. Joy	0	0	1	2	14	27
11. Triumph	0	0	0	1	0	2

For each of these phrases there seems to be a critical speed beyond which changes occur at a slower rate. For Phrase A, this is at 120, the slower tempos not showing much divergence in the distribution of responses. For Phrase B, it is at 120, the faster speeds not showing so much change. When Phrase C is played at 100 or slower, the results do not vary appreciably. Phrase D, played faster than 138, is not much happier, although the triumph replies continue to increase. For Phrase E, there is not much change above 120.

It is evident from the data that a faster speed tends to make the music happier (or more joyful), while a slower speed makes it sadder (or more of a lamentation). For the most part this result appears clearly in the relative strength of the *X* and *Y* responses. At other times it is to be seen in the subheadings, in the changes for the categories melancholy, lamentation, gaiety, and joy. There is only one slight inversion in the *X* and *Y* votes (Phrase E at speeds 138 and 160), and even here the general principle is manifested in the gains for gaiety and joy at the faster speed. The only other partial

exception to the finding is Phrase C for the speeds 60 and 80. Here there is practically no difference in the two patterns of response.

Summary.—Five short musical phrases, three happy in mood and two sad, were played, each at six different metronome speeds, to 88 observers. The general principle that fast tempo tends to make music happy while slow tempo has the opposite effect, was abundantly illustrated in the data.

(Manuscript received June 17, 1940)

INFLUENCE OF LINE WIDTH ON EYE MOVEMENTS [1]

BY DONALD G. PATERSON AND MILES A. TINKER

University of Minnesota

In our investigations of the influence of various typographical arrangements on speed of reading we completed eleven studies of line width.[2] In general the results indicate that line widths can be varied to a surprising degree without any appreciable adverse effect on speed. That is, a definite range of optimal line widths exists. However, lines shorter than or longer than the optimal range definitely retard reading speed.

Having established, by reading performance tests, the optimal range of line widths for ten point type set solid, we undertook two eye-movement studies designed to reveal the specific patterns of eye movements involved in reading optimal versus non-optimal line widths. The eye movements were photographed by the Minnesota eye-movement camera.

In the first study, each of 20 college students read 10 paragraphs from the Chapman-Cook Speed of Reading Test, Form A, set in an optimal line width (19 picas) and 10 different paragraphs from Form B of the same test set in an excessively short line width (9 picas). In the second study, each of 20 additional college students read ten paragraphs from Form A in the 19 pica width and 10 paragraphs from Form B in the excessively long line width of 43 picas. Figure 1 shows samples of the reading material employed in both studies.

The photographic records yielded for each study the following five analytical measures: number of fixations, number of words per fixation, duration of pauses, perception time, and number of regressions.

[1] The writers are grateful to the University of Minnesota Graduate School for research grant to finance this study.

[2] Detailed results of these and other studies will be published shortly by Harper & Brothers under the title, *How to Make Type Readable.*

vestigation, that the periodic-continuous sequence would operate to *reduce* the number of responses during extinction. Such an effect, if obtained, would bear a specious resemblance to Dunlap's findings with respect to the breaking of undesirable habits (1) and, at the same time, be consistent with the view that the rat stops responding sooner because he has come to expect a reward for *every* lever-depression. Although it is unlikely that the results herein reported will escape the generality of 'expectancy' interpretation, it is quite certain that they do *not* provide us with a new technique for the quicker removal of well-entrenched habits—for subtracting from the reflex reserve by the simple operation of reinforcement.

(Manuscript received May 29, 1940)

References

1. Dunlap, K., *Habits: their making and unmaking*, N. Y., Liveright, 1932, x + 326.
2. Ellson, D. G., The concept of reflex reserve, *Psychol. Rev.*, 1939, **46**, 566–575.
3. Goulden, C. H., *Methods of statistical analysis*, N. Y., Jno. Wiley & Sons, 1939, p. 267.
4. Keller, F. S., Some experimental operations affecting drive constancy, *J. comp. Psychol.*, 1939, **28**, 63–72.
5. Skinner, B. F., *The behavior of organisms*, N. Y., Appleton-Century, 1938, pp. 457.

SPEED AS A DETERMINER OF MUSICAL MOOD

BY MELVIN G. RIGG

Oklahoma A. and M. College

From previous experiments it has appeared that tempo has more influence upon the emotional suggestiveness of music than any other single factor.[1] This feature would therefore seem to deserve a more exhaustive study than it has hitherto received. The present paper reports an experiment dealing with speed alone.

The stimuli consisted of five musical phrases played on the piano. The phrases were harmonized and were four measures in length.[2] Each was presented at six different speeds, the tempo being rigidly controlled by a metronome.[3] The phrases are designated as A, B, C, D, and E with a number following to indicate the speed (the number of quarter notes to the minute).

The order of playing was as follows:

(read down)		
E 120	E 138	E 160
D 120	D 138	D 152[4]
C 120	C 138	C 160
B 120	B 138	B 160
A 120	A 138	A 160
E 100	E 80	E 60
D 100	D 80	D 60
C 100	C 80	C 60
B 100	B 80	B 60
A 100	A 80	A 60

The observers in the experiment were 88 college students enrolled in psychology classes. They represented varying de-

[1] Rigg, M. G., What features of a musical phrase have emotional suggestiveness? *Pub. Soc. Sci. Res. Coun. Oklahoma A. and M. Coll.*, 1939, 1, 38 pp.

[2] These phrases were used in the study previously mentioned, *op. cit.* They may be found in musical notation in the appendix, designated as A 1, B 1, C 1, D 1, and E 1.

[3] It was at first feared that the sound of the metronome would have an influence upon the emotional effect of the music, but preliminary experimentation revealed that such was not the case.

[4] It was beyond the technical ability of the pianist always to give a perfect rendition of this phrase at 160.

grees of musical training, but it seems unlikely that musical training influences judgments concerned with the emotional effects of music.[5]

Each auditor was supplied with a check list to assist him in making his replies. From previous studies, the most valid discrimination with regard to the emotional character of music appeared to be between the happy and the sad. The observers were consequently asked to designate whether the phrase was primarily sad-serious (indicated as *S*) or pleasant-happy (indicated as *Y*). After this judgment the auditors were to attempt a finer discrimination by classifying the phrase under subheadings as follows:

X Serious-sad
- 1 Solemnity
- 2 Sorrowful longing
- 3 Melancholy
- 4 Lamentation
- 5 Agitation

Y Pleasant-happy
- 6 Hopeful longing
- 7 Love
- 8 Revery
- 9 Gaiety
- 10 Joy
- 11 Triumph [6]

It was explained to the observers that melancholy was to mean something less intensely sad than lamentation, and gaiety, something less intense than joy.

The data for Phrase A are presented in Table I. The results for the various renditions of this phrase are given not in the order of performance, but in that of increasing speed. When this phrase was played 60 quarter notes to the minute, the *X* votes were unanimous, and 37 of these were specifically

[5] Hevner, K., The affective character of the major and minor modes in music, *Amer. J. Psychol.*, 1935, **47**, 103–118. Rigg, M. G., An experiment to determine how accurately college students can interpret the intended meanings of musical compositions, *J. exper. Psychol.*, 1937, **21**, 223–229.

[6] This list is identical with the one used in the previous investigation. See first footnote.

TABLE I

PHRASE A

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	88	86	86	86	80	75
1. Solemnity	24	22	20	22	11	12
2. Sorrowful Longing	7	7	9	10	4	3
3. Melancholy	15	14	12	23	17	8
4. Lamentation	37	34	32	19	21	11
5. Agitation	5	9	13	12	27	41
Y Pleasant-Happy	0	1	2	2	8	13
6. Hopeful Longing	0	0	0	2	5	7
7. Love	0	1	0	0	0	0
8. Revery	0	0	0	0	1	4
9. Gaiety	0	0	2	0	1	0
10. Joy	0	0	0	0	0	0
11. Triumph	0	0	0	0	1	2

for lamentation. The *X* responses remain at 86 for the three next faster speeds, but the lamentation votes go down successively to 34, 32, and 19. It is safe to conclude therefore that the phrase is progressively less intense in its sadness. At the two faster speeds (138 and 160) the *X* responses go down, first to 80 and then to 75. Even at the fastest speed the phrase is still sad, but the judgment is no longer unanimous and there is a tendency for the music to suggest agitation.

The data for Phrase B are presented in Table II. This

TABLE II

PHRASE B[7]

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	32	24	15	9	6	5
1. Solemnity	3	2	1	0	0	0
2. Sorrowful Longing	5	7	3	1	1	0
3. Melancholy	3	1	2	0	0	1
4. Lamentation	6	3	1	2	1	1
5. Agitation	15	11	8	6	4	3
Y Pleasant-Happy	54	63	73	79	82	83
6. Hopeful Longing	23	19	9	4	2	2
7. Love	2	7	5	3	1	3
8. Revery	12	13	11	6	7	10
9. Gaiety	14	17	31	40	31	20
10. Joy	2	5	13	22	34	42
11. Triumph	1	2	4	4	5	6

[7] Since occasionally an auditor would mark a phrase *X* or *Y* without making the second discrimination, the total for the subheadings may be slightly less than the number for the main heading.

phrase shows a different picture, since it is primarily pleasant-happy in character. For every increase of speed the *Y* votes become more numerous. Between the two fastest speeds this gain is only 1; nevertheless, the joy responses advance from 34 to 42.

TABLE III

PHRASE C

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	81	81	79	67	58	43
1. Solemnity	6	10	7	11	10	6
2. Sorrowful Longing	31	26	25	24	17	11
3. Melancholy	26	30	29	19	12	4
4. Lamentation	15	15	17	11	8	6
5. Agitation	3	0	1	2	11	16
Y Pleasant-Happy	7	7	9	21	30	45
6. Hopeful Longing	4	5	2	9	13	22
7. Love	1	2	4	5	7	7
8. Revery	2	0	3	4	7	9
9. Gaiety	0	0	0	2	2	4
10. Joy	0	0	0	1	0	2
11. Triumph	0	0	0	0	1	1

The data for Phrase C are presented in Table III. This phrase is definitely sad except at the fastest tempo. The changes between the two slowest speeds appear inconsequential, but the *Y* votes show a steady gain at the fastest speeds.

The data for Phrase D are presented in Table IV. At slow speed this phrase is sad; the faster it is played, the happier it becomes. At the fastest tempo it suggests triumph.

TABLE IV

PHRASE D[7]

Metronome Speeds	60	80	100	120	138	152
X Serious-Sad	50	44	19	13	4	1
1. Solemnity	13	16	8	6	1	0
2. Sorrowful Longing	22	15	7	4	0	0
3. Melancholy	4	5	2	2	1	0
4. Lamentation	8	6	1	1	1	0
5. Agitation	2	2	1	0	1	1
Y Pleasant-Happy	38	44	68	75	84	87
6. Hopeful Longing	18	15	27	17	8	4
7. Love	1	6	5	2	2	2
8. Revery	4	7	11	15	7	4
9. Gaiety	0	1	3	11	5	7
10. Joy	2	3	4	4	17	11
11. Triumph	13	12	17	26	45	59

The data for Phrase E are presented in Table V. Here there is one inversion, since for speed 160 there are only 85 *Y* responses as against 87 for speed 138. However, gains are registered in both the categories of joy and gaiety for the faster of these tempos.

TABLE V

PHRASE E

Metronome Speeds	60	80	100	120	138	160
X Serious-Sad	41	33	19	5	1	3
1. Solemnity	11	9	5	2	0	0
2. Sorrowful Longing	16	15	9	2	0	1
3. Melancholy	8	6	4	0	0	1
4. Lamentation	4	3	1	1	1	0
5. Agitation	2	0	0	0	0	1
Y Pleasant-Happy	47	55	69	83	87	85
6. Hopeful Longing	16	27	30	29	14	11
7. Love	18	15	26	18	16	4
8. Revery	13	13	11	28	18	8
9. Gaiety	0	0	1	5	25	33
10. Joy	0	0	1	2	14	27
11. Triumph	0	0	0	1	0	2

For each of these phrases there seems to be a critical speed beyond which changes occur at a slower rate. For Phrase A, this is at 120, the slower tempos not showing much divergence in the distribution of responses. For Phrase B, it is at 120, the faster speeds not showing so much change. When Phrase C is played at 100 or slower, the results do not vary appreciably. Phrase D, played faster than 138, is not much happier, although the triumph replies continue to increase. For Phrase E, there is not much change above 120.

It is evident from the data that a faster speed tends to make the music happier (or more joyful), while a slower speed makes it sadder (or more of a lamentation). For the most part this result appears clearly in the relative strength of the *X* and *Y* responses. At other times it is to be seen in the subheadings, in the changes for the categories melancholy, lamentation, gaiety, and joy. There is only one slight inversion in the *X* and *Y* votes (Phrase E at speeds 138 and 160), and even here the general principle is manifested in the gains for gaiety and joy at the faster speed. The only other partial

exception to the finding is Phrase C for the speeds 60 and 80. Here there is practically no difference in the two patterns of response.

Summary.—Five short musical phrases, three happy in mood and two sad, were played, each at six different metronome speeds, to 88 observers. The general principle that fast tempo tends to make music happy while slow tempo has the opposite effect, was abundantly illustrated in the data.

(Manuscript received June 17, 1940)

INFLUENCE OF LINE WIDTH ON EYE MOVEMENTS [1]

BY DONALD G. PATERSON AND MILES A. TINKER

University of Minnesota

In our investigations of the influence of various typographical arrangements on speed of reading we completed eleven studies of line width.[2] In general the results indicate that line widths can be varied to a surprising degree without any appreciable adverse effect on speed. That is, a definite range of optimal line widths exists. However, lines shorter than or longer than the optimal range definitely retard reading speed.

Having established, by reading performance tests, the optimal range of line widths for ten point type set solid, we undertook two eye-movement studies designed to reveal the specific patterns of eye movements involved in reading optimal versus non-optimal line widths. The eye movements were photographed by the Minnesota eye-movement camera.

In the first study, each of 20 college students read 10 paragraphs from the Chapman-Cook Speed of Reading Test, Form A, set in an optimal line width (19 picas) and 10 different paragraphs from Form B of the same test set in an excessively short line width (9 picas). In the second study, each of 20 additional college students read ten paragraphs from Form A in the 19 pica width and 10 paragraphs from Form B in the excessively long line width of 43 picas. Figure 1 shows samples of the reading material employed in both studies.

The photographic records yielded for each study the following five analytical measures: number of fixations, number of words per fixation, duration of pauses, perception time, and number of regressions.

[1] The writers are grateful to the University of Minnesota Graduate School for research grant to finance this study.

[2] Detailed results of these and other studies will be published shortly by Harper & Brothers under the title, *How to Make Type Readable.*

Standard: 19 picas

11. When my mother saw the marks of muddy shoes on the floor, and all over the nice clean beds, she was surprised to see how careful the children had been. 12. When the little boy next door had both

Short line: 9 picas

11. Frank had been expecting a letter from his brother for several days; so as soon as he found it on the kitchen table he ate it as quickly as possible. 12. A cer-

Long line: 43 picas

11. Frank had been expecting a letter from his brother for several days; so as soon as he found it on the kitchen table he ate it as quickly as possible. 12. A certain doctor living in a city near here always has a very serious expres-

FIG. 1. Showing samples of the three line widths.

19 Picas versus 9 Picas

The results for the first study (19 versus 9 pica line widths) are shown in Table I.

TABLE I

Mean Eye-Movement Measures for 20 College Subjects
Reading Text Set in 19 Pica and 9 Pica Line Widths

Note: Pause duration and perception time are reported in seconds. Ten paragraphs from the Chapman-Cook Speed of Reading Test, Form A, were set in a 19 pica line width and 10 paragraphs from Form B of the same test were set in a 9 pica line width. All paragraphs were printed as follows: Scotch Roman, lower case, 10 point, set solid on egg-shell paper stock.

Line Width	Fixation Frequency	Words per Fixation	Pause Duration	Perception Time	Regression Frequency
19 picas	193.5	1.6	0.22	42.8	26.2
9 picas	223.9	1.4	0.24	53.4	22.5
Differences *	+30.5	−0.2	+0.02	+10.6	−3.7
Percent Diff.	+15.7	−14.3	+8.1	+24.7	−14.2

* All differences are statistically significant except for regression frequency.

Scrutiny of these results shows a marked increase (15.7 percent) in the average number of fixations in reading the 9 pica line width in comparison with the 19 pica optimum. On the average fewer words were perceived per fixation when the short line was read. Curiously enough, more time in reading the short line was spent on each fixation (pause duration percentage difference equals 8.1). Also for the short line there is a striking increase (24.7 percent) in the total perception time. The only apparent advantageous characteristic of the short line width is found in the fact that it was read with fewer regressions (14.2 percent decrease). When the significance of the difference in the means of correlated measures was computed and the *t* values ascertained it was found that all *t* values were beyond the 1 percent level except that for regressions which was less than the 10 percent level.

One may characterize the oculomotor patterns in reading an excessively short line by saying that the number of fixations is increased, the span of perception is decreased, the mean duration of the fixations is increased, total perception time is greatly increased but the number of regressions is

approximately the same. It is seen, therefore, that characteristic differences in oculomotor patterns appear when reading an excessively short line width in comparison with the optimum. As may be observed in Fig. 1, each short line contains about 4 or 5 words whereas the 19 pica line width contains about 10 words. There were 3.5 fixations per line in the former case and 6.4 fixations per line in the latter. This means that the eye seems to be unable to maintain as large a perceptual span (words per fixation) in reading the short line. The net result is more fixations, a longer time on each fixation, with a consequent great increase in total perception time. It is obvious that the short line does not involve distortions of word forms nor reduction in visibility of the printed characters. The only remaining factor that could account for the decreased efficiency in reading the short line would appear to be the difficulty in making maximum use of peripheral vision in the horizontal direction. In such a short line, the eye is forced after only 3 or 4 fixations to swing back to the next line. Thus it is impossible to capitalize efficiently on horizontal peripheral vision since the sequence of eye movements is interrupted so frequently. That is, the ability to increase the span of perception and to decrease pause duration which ordinarily occurs in reading the second half of an optimal line width is seriously interfered with.

19 Picas versus 43 Picas

The results for the second study (19 versus 43 picas) are shown in Table II. At the outset it should be stated that all of the differences in Table II are statistically significant (all t values are beyond the 1 percent level).

Examination of the detailed results in Table II shows that fixation frequency in reading the long line is definitely increased, whereas the average number of words per fixation is decreased. Furthermore, the average time devoted to fixations (pause duration) is slightly increased and, of course, total perception time is increased as well. A most striking difference occurs with respect to the number of regressions. This amounts to 56.7 percent increase. Inspection of the

photographic records reveals that practically every line in the 43 pica test-copy involves a regression. These come chiefly at the beginning of each line.

TABLE II

MEAN EYE-MOVEMENT MEASURES FOR 20 COLLEGE SUBJECTS
Reading Text Set in 19 Pica and 43 Pica Line Widths

Note: Pause duration and perception time are reported in seconds. Ten paragraphs from the Chapman-Cook Speed of Reading Test, Form A, were set in a 19 pica line width and 10 paragraphs from Form B of the same test were set in a 43 pica line width. All paragraphs were printed as follows: Scotch Roman, lower case, 10 point, set solid on egg-shell paper stock.

Line Width	Fixation Frequency	Words per Fixation	Pause Duration	Perception Time	Regression Frequency
19 picas	190.1	1.7	0.22	41.3	27.4
43 picas	204.9	1.5	0.23	47.0	42.9
Differences *	+14.8	− 0.2	+0.01	+ 5.7	+15.5
Percent Diff.	+ 7.8	−10.6	+ 3.8	+13.8	+56.7

* All differences are statistically significant.

We interpret the results to mean that an excessively long line width gives rise to a major difficulty in swinging back to the beginning of successive lines. In other words, the eyes are inaccurate in relocating the beginning of each new line. It is entirely possible that this difficulty so upsets the reading process that the reëstablishment of the most efficient oculomotor patterns in reading each line is impossible.

SUMMARY

1. In a series of eleven reading performance experiments the writers demonstrated that excessively short lines (9 picas) and excessively long lines (43 picas) are both read much more slowly than lines of moderate width (19 picas).

2. Eye-movement photographs of two new groups of 20 subjects each indicate the precise changes in oculomotor patterns involved in reading excessively short and long line widths.

3. It appears that the decreased efficiency with which a very short line is read is due to the reader's inability to make maximum use of horizontal peripheral cues.

4. Analytical records derived from the reading of extremely long line widths reveal that the task of locating the beginning of successive lines of print is accomplished only with great difficulty. This is shown by the striking increase in regression frequency. It is as if the eyes, in making the return sweep, are forced to spend an undue amount of time in searching for the next line. Presumably this difficulty tends to upset the smooth functioning of oculomotor habits.

(Manuscript received June 17, 1940)

A NEW APPARATUS FOR THE LURIA EXPERIMENT

BY RUTH A. BOBBITT

Mills College, California

The apparatus described below is designed to fulfill the same basic purposes as that devised by A. R. Luria,[1] namely the study of affect on the basis of certain motor manifestations. The typical experimental procedure employed is that of the 'free-association' experiment. The subject holds each of the index fingers upon a specially constructed key, and he is instructed to depress the right-hand key at the time that he utters his verbal response. The apparatus records verbal stimulus and response, as well as the position and movement of each index finger throughout the experimental session.

The present apparatus has the advantages of (1) permitting the subject and the experimenter to be situated in different rooms, and (2) of furnishing a record not merely of the onset of the verbal response and its reaction-time, but of the form or structure of this response.

Submitted are photographs of the apparatus as it is set up in two rooms; one from which the experimenter conducts the experiment (Fig. 1), and the other in which the subject is seated (Fig. 2).

(1) *Verbal Communication System.*—A carbon microphone (*A*, Fig. 1), on an adjustable telescopic rod, is set just in front of the subject. The experimenter is equipped with a portable microphone (*A*, Fig. 2). Both the subject and the experimenter wear headphones (*B*, Figs. 1 and 2). As the subject speaks into his microphone, his voice sets up electrical impulses that pass into a four-stage amplifier (*C*, Fig. 1). The amplified voice-currents are taken from the output stage of this amplifier and passed through a rectifier-circuit of the diode type. They are transformed from alternating to direct-

[1] Luria, A. R., *The Nature of Human Conflicts* (translated from the Russian and edited by W. Horsely Gantt), New York: Liveright, 1932, 431 pp.

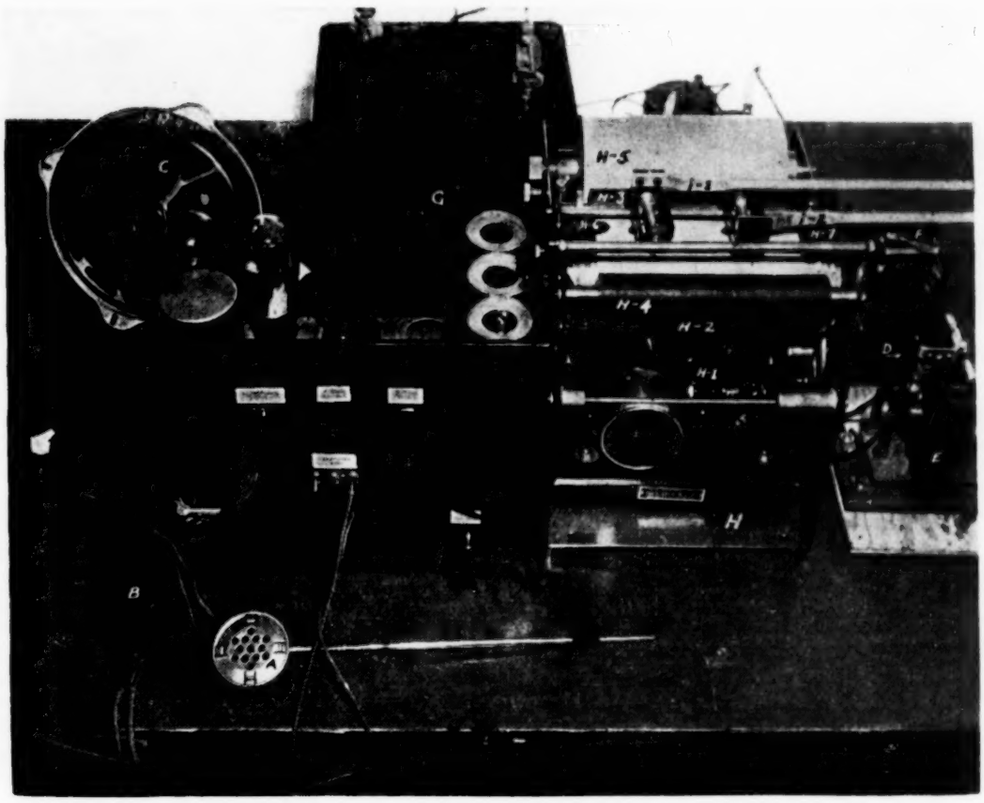


FIG. 1. Apparatus as set up in the experimenter's room.

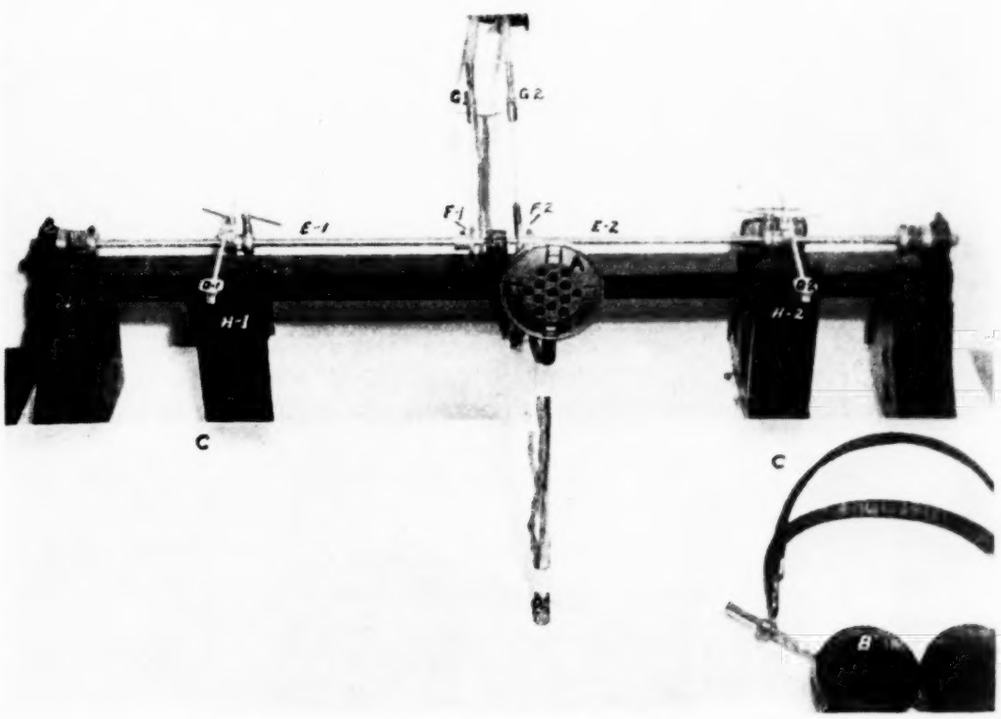


FIG. 2. Apparatus as set up in the subject's room.

current impulses at radio frequency. These direct-current impulses operate a sensitive 5000-ohm voice-controlled relay (*D*, Fig. 1) which, in turn, activates an ordinary power-relay (*E*, Fig. 1). (The power-relay has contacts heavy enough to carry the current that is required for the recording-stylus magnet (*F*, Fig. 1) (*vide infra*).) A six-volt storage battery (*G*, Fig. 1) supplies current, via the power-relay contacts, to the recording-stylus magnet. The sensitivity of this arrangement is entirely sufficient. (Indeed one finds it necessary to turn down the 'gain-control' during actual use. When the applied sound is intense, the springy extension-arm of the stylus-magnet stays over until the reverberations of the sound die down.)

The communication-circuit between the subject and the experimenter makes use of the voice-amplifier to give the desired sound-intensity in the headphones. The microphones are operated by a separate set of dry cells in order that the electrical surge caused by the stylus-magnet will not distort the voice in the earphones. The complete communication equipment is arranged as follows: the amplifier and diode-rectifier, together with the power-supply, are built on a metal chassis. The chassis has a front panel with jacks for microphones, headphones, relays, and connections for regulating sensitivity and headphone current. The two relays are assembled on a separate board for convenience and simplicity. The stylus-magnet coil and stylus are attached to a platform on one side of the paper-driving device (*H*, Fig. 1) (*vide infra*). The voice-recordings are made on one extreme edge of the stylographic paper.

(2) *Device for Recording Finger-Movements.*—The subject is seated at a table on which rests the equipment for recording his finger-movements. All of this apparatus is mounted on a portable board, 22 × 29 inches. The subject's forearms rest on a board (*C*, Fig. 2), so inclined that his wrists are 3½ in above the table surface. The index finger of each hand rests on a key which is attached to a horizontal lever-arm. The distance between the lever-arms may be adjusted to the convenience of the subject. Each of these arms is fastened at

right angle to a horizontal steel rod, 12 in in length and 5/16 in in diameter (*E*1, *E*2, Fig. 2). Each rod is so pivoted to vertical uprights that it may be independently rotated by the slightest movement of the attached key. The two vertical uprights, into each of which the inside end of one of the rods is pivoted (*F*1, *F*2, Fig. 2), are swung backward or forward as controlled by the rotation of the rods. At the top, each upright also has a socket (*G*1, *G*2, Fig. 2) in which rests one end of an aluminum tube, 36 in long and 5/16 in in external diameter. These tubes extend to the paper-driving device in the experimenter's room. There they are attached to stylus-carriages (*I*1, *I*2, Fig. 1) which move freely along a track just above the recording paper.

As the subject depresses one of the keys, the attached steel rod is rotated. It draws the vertical upright, the aluminum tube, and the stylus at the other end of the tube, toward the subject. When the key is released the process is reversed. Thus a curve showing the variations of pressure is obtained.

The right-hand lever-key has sufficient tension to bring it back to its normal position after it has been depressed by the subject. Various spring tensions may be put on the left-hand lever, and the length of the lever-arm itself may also be adjusted to variations in tremor-amplitude. The amplitude of movement of each hand is further restricted by wooden blocks (*H*1, *H*2, Fig. 2) placed at such a distance below the lever-arms as to prevent the keys from exceeding a ¾-inch span. The entire lever-system, with the exception of the keys, is enclosed in a wooden housing during the experiment.

(3) *Paper-Driving Device.*—The paper-driving device (*H*, Fig. 1) is small, compact, and portable. Its design embodies a sheet-brass housing, 10 × 9 × 8 in, of which the top and two ends are left open. A 110 volt, A.C., 1/75 H. P., Bodine synchronous motor (*H*1, Fig. 1) is geared so that it causes a grooved brass drum (*H*2, Fig. 1) of 10 in circumference to make one revolution every 20 seconds. The motor is mounted on a wooden soleplate which, in turn, is set on rubber on the housing floor. In this way, motor vibration is decreased to a

minimum. On each side of the brass drum a rubber-encased roller (*H*3, *H*4, Fig. 1), 1 1/8 in in diameter, is held against the drum by means of an equated spring tension. A roll of stylographic recording-paper (*H*5, Fig. 1) is placed securely on a brass spool. The spool revolves as the drum and rollers draw the paper from it at a constant speed of ½ in per second. The paper is held taut and straight as it is drawn between one roller and the drum, over the drum, and out between the other roller and the drum.

Four stylus attachments are mounted on the paper-driving device as follows:

(*a*) A time-recording stylus (*H*6, Fig. 1) is operated by means of 20 evenly spaced cogs in one end of the drum. In accord with the revolution speed of the motor, the stylus marks seconds at every half inch on the extreme right edge of the paper as it is drawn over the drum.

(*b*) The voice-reaction stylus (*H*7, Fig. 1), as noted above, extends from the electro-magnet, mounted on the left side of the housing, to the other edge of the stylographic paper.

(*c*) To each of the aluminum tubes (*I*1, *I*2, Fig. 1), manually operated by the subject, is fastened a light carriage with specially ground pulleys. From the bottom of each of these carriages extends a fine steel-tipped stylus. Just above the surface of the recording-drum are fastened two rods, 8 in in length, 3/16 in in diameter, and 1 ¼ in apart, which serve as the track along which the two stylus-carriages are drawn.

The complete apparatus is relatively inexpensive to construct. Its design is flexible so that minor adjustments may be made as they are found necessary by actual experimentation.

(Manuscript received June 19, 1940)